Trust and Reciprocity:
Implications of Game Triads and Social Contexts

By James C. Cox

University of Arizona
jcox@eller.arizona.edu

Revised May 2003

Abstract

This paper uses a triadic experimental design to conduct trust and reciprocity experiments. The experiments with game triads discriminate between transfers resulting from trust or reciprocity and transfers resulting from (intentions-unconditional) other-regarding preferences. Alternative treatments vary the environment of the experiments from a weak to a strong social context. The observations in two social contexts of decisions motivated by other-regarding preferences, trust, and/or reciprocity are central to obtaining empirical information that can guide the process of formulating a theory of utility that can increase the empirical validity of game theory. Data from experiments with the triadic design are used to evaluate recent extensions of theory that incorporate concern for fairness of monetary payoffs and perceptions of others' intentions into agents' utilities.

Keywords: experimental economics, game theory, trust, reciprocity, altruism

JEL Classification: C70, C91, D63, D64

Trust and Reciprocity:
Implications of Game Triads and Social Contexts*

1. Introduction

I use the triadic experimental design previously reported in Cox (2002, forthcoming). This design makes it possible to discriminate between transfers resulting from trust or reciprocity and transfers resulting from (intentions-unconditional) other-regarding preferences. Being able to make the discriminations between the implications of such other-regarding preferences and trust or reciprocity is important to obtaining the empirical information that can guide the process of formulating a theory of utility that can increase the empirical validity of game theory. In addition to identifying transfers with alternative motivations, the data are also used to evaluate four recently-published extensions of theory.

The present paper varies the social context, in which the triadic game structure is implemented, from a "weak" (or one-task) to a "strong" (or two-task) social context. Use of the strong social context makes it possible to measure the effects of trust, reciprocity, and other-regarding preferences in an environment in which there will be future social interaction but no repeated games. This introduces into the experiment (in a minimal way) a feature of everyday life in cities. Much interaction in urban and rural environments involves repeated games, such as interactions among family members and co-workers, but much social interaction in large cities involves play of one shot games in rich social contexts. Examples which illustrate this point are provided by automobile and bus driver interactions in Tucson, Arizona and Oxford, England. I will provide some stylized facts about those driver interactions after first defining trust and reciprocity.

2. Definitions and Stylized Facts

Interpretations of data in this paper will be based on the following definitions. Preferences over another individual's money payoff, in addition to one's own money payoff, will be referred to as "other-regarding preferences." Such preferences can be altruistic or malevolent. Let y^j denote the amount of income of agent *j*, and assume that agent *k*'s preferences can be represented by a utility function. Then agent *k* has other-regarding preferences for the income of agent *j* if his/her utility function, $u^k(y^k, y^j)$ is *not* a constant function of y^j. It is important to distinguish other-regarding preferences from motivations involving trust and reciprocity.

The concepts of trust and positive reciprocity used in this paper are defined as follows. Agent 1 undertakes an action that exhibits "trust" if the chosen action: (a) creates a monetary gain that could be shared with Agent 2; and (b) exposes Agent 1 to the risk of a loss of utility if Agent 2 defects and appropriates too much of the monetary gain. Agent 2 undertakes an action that exhibits "positive reciprocity" if the chosen action: (a) gives Agent 1 a monetary gain; and (b) is undertaken instead of an available alternative action that would produce outcomes preferred by Agent 2 in the absence of the trusting action by Agent 1. Note that the definition of reciprocity incorporates a possible dependence of preferences over outcomes upon the process that generated those outcomes and the perceived intentions of others. This would seem to be a necessary condition for rational agents to undertake actions involving trust and reciprocity. Thus, Agent 1 can rationally undertake a trusting action if he/she believes that this choice may trigger a social norm in Agent 2 that causes him/her not to defect. Alternatively, Agent 1 can rationally undertake a trusting action if he believes that Agent 2 has altruistic other-regarding preferences. The experiment design for game triads explained in section 5 makes it possible to discriminate between the implications of other-regarding preferences and trust or reciprocity. This discrimination is central to developing alternatives to the completely selfish model of agents' utilities.

The possible dependence of the incidence of other-regarding, trusting, and reciprocating actions on the social context of the decision environment can be appreciated from considering the following commonplace example of automobile driving behavior in Tucson, Arizona. Tucson and adjacent communities in Pima County comprise a medium-size metropolitan area with population of about 800,000. Because there is only one freeway on the edge of the city, much of the commuter traffic is carried on city streets. The heavy traffic produces frequent interactions between drivers in traffic lanes who own the (both legally- and socially-defined) right-of-way and other drivers wanting to enter the traffic lanes from parking lots and driveways. Each interaction between a pair of drivers, one with and the other without ownership of the lane right-of-way, is (for any practical purpose) a one-shot game *in the sense that* the probability that this pair of drivers will ever encounter each other again is close to 0. Thus it is not possible to develop a personal reputation with any other individual driver that can provide benefit or cost in the future. But the one-shot games are played in a social context that has some observable implications. My observations of play of this right-of-way game are as follows.

Many drivers refuse to let the other driver in during an individual encounter. In a substantial minority of the encounters, the driver owning the right-of-way delays his or her trip to let the paired driver into the traffic lane. In response, most but not all of the drivers receiving the courtesy wave and smile to convey their gratitude. But because of the social context of play, such an interaction between a pair of drivers is not always the end of the actions triggered by the courtesy of the original driver who gave the right-of way to another. In some cases, one can see the recipient of the original courtesy extend a similar one to a third party. My personal experience as a participant is as follows. If another driver extends me the courtesy of letting me into the street, the probability is essentially 1 that I will extend the same courtesy to the next driver I encounter on that trip.[1] That such "second task" behavior can have implications for the first encounter between two paired drivers can be understood as follows. Suppose that some drivers prefer that other drivers be courteous. Also suppose that some of the drivers with this preference anticipate that courtesy can be contagious. Then these drivers will have a greater motivation to extend the first courtesy themselves than they would if their one-shot interaction with a specific other driver were not embedded in a social context of numerous one-shot interactions between pairs of drivers. Thus the richness of the social context of play of one-shot games may be a significant determinant of behavior.

During his journey from Oxford, England to Cambridge, Massachusetts where he would present his paper and discuss an earlier draft of this paper, Jim Engle-Warnick made some informal observations that supplement mine. While riding the bus from Oxford to Heathrow Airport, Dr. Engle-Warnick observed interactions between the bus driver and pedestrians, cyclists, and other drivers. He counted 12 courtesies extended by the bus driver to others, and six courtesies extended by others to the bus driver. There were eight (directly-reciprocal) waves of acknowledgement by recipients of courtesies. Even more interestingly, there were three observations of recipients of courtesies in turn extending new courtesies to third parties (indirect reciprocity) within the short period of time before the original recipients disappeared from view. Upon being asked at the end of the trip why he was courteous to others on the road, the bus driver replied: "Because other people will be more likely to do it."

The interactions among drivers, cyclists, etc. involve patterns of behavior similar to those studied in laboratory fairness games. Thus a driver may incur a cost to extend a courtesy to another because of other-regarding preferences or because of trust that the other person may reciprocate. The individual receiving the courtesy may wave and smile to show positive reciprocity. Furthermore, a driver receiving a courtesy may be more likely to subsequently extend a similar courtesy to another, thus exhibiting indirect reciprocity. And, of course, there are examples of discourteous driving behavior with associated patterns of negatively reciprocal responses such as honking, curing, etc. These interactions are, arguably, one-shot games because the probability that the same pair of drivers will encounter each other more than once is close to zero in an urban environment. But the one-shot games are played in a social context with the following salient characteristic: while extending a courtesy or responding to one, a driver knows that other interactions lie ahead but does not know exactly what those interaction (or games) will turn out to be. The design of our rich social context experiment is intended to incorporate this salient characteristic of everyday social interaction into the laboratory in a simple stylized way by informing the subjects that there will be a subsequent task with possible monetary payoff but telling them nothing specific about the task.

The experiment design explained in section 5 varies the environment from a weak to a strong social context. The experiment involves game triads that include the "investment" game introduced by Berg, Dickhaut, and McCabe (1995) and later used by several other authors.

3. The Investment Game

Berg, Dickhaut, and McCabe (hereafter BDMc) implemented the investment game by dividing the subjects into two groups, the Room A group and the Room B group. Every subject was given an endowment of ten $1 bills. Room B subjects were instructed to keep their $10 endowments. Room A subjects were informed that they could keep all of their $10 endowments or transfer any integer amount to a paired subject in Room B. Any amount transferred by a Room A subject was multiplied by 3 by the experimenter before being delivered to a Room B subject. Subsequently, Room B subjects were given the opportunity to return none, part, or all of the amount received to the paired subject in Room A. The experiments used a double-blind payoff protocol in which individual subjects' responses were anonymous to both other subjects and the experimenter.

If one assumes there are no other-regarding preferences, then standard game theory predicts that: (i) Room B subjects will keep all the money they receive because Room B subjects prefer more money to less; and (ii) knowing this, Room A subjects will not transfer any money. This "completely-selfish" subgame perfect equilibrium leaves each pair of subjects with $20, whereas it could have ended up with as much as $40. Thus this allocation is Pareto inferior to some alternative feasible allocations.

Results from investment-game experiments reported by BDMc were that the average amount transferred by Room A subjects was $5.16 and the average amount returned by Room B subjects was $4.66. When data from this experiment were provided to subjects in a subsequent experiment (the "social history" treatment), the average amount transferred by Room A subjects was $5.36 and the average amount returned was $6.46. There was large variability across subjects in the amounts transferred and returned. The experiments reported by BDMc used a "double blind" protocol in which subjects' responses were anonymous to other subjects and the experimenters.

Data from investment game experiments support the following interpretation. A Room A subject may be willing to transfer money to an unknown Room B person if he or she *trusts* that some of the tripled amount transferred will be returned. Further, a Room B subject may be willing to return part of the tripled amount transferred if he or she practices *positive reciprocity.* But a Room A subject may be willing to make a transfer to a paired subject in Room B even if there is no opportunity for the latter to return anything. Data from the investment game do not allow one to distinguish between the level of transfers resulting from trust and that resulting from altruistic other-regarding preferences. Similarly, investment game data do not discriminate between transfers resulting from reciprocity and (intentions-unconditional) other-regarding preferences because a Room B subject may be willing to transfer money to a paired subject in Room A even if that paired subject did not give him/her any money. In order to make these distinctions, one needs a more elaborate experimental design.

4. The Triadic Experimental Design

The experiment involves three treatments implemented in an across-subjects design. Treatment A is the investment game. Each individual in the second-mover group is credited with a $10 endowment. Each individual in the first-mover group is credited with a $10 endowment and given the task of deciding whether she wants to transfer to a paired individual in the other group none, some, or all of her $10. Any amounts transferred are tripled by the experimenter. Then each individual in the second-mover group is given the task of deciding whether he wants to return some, all, or none of the tripled number of certificates he received to the paired individual in the other group.

Treatment B is a dictator game that differs from treatment A only in that the individuals in the “second-mover” group do not have a decision to make; thus they do not have an opportunity to return any tokens that they receive.

Treatment C involves a decision task that differs from Treatment A as follows. First, the “first movers” do not have a decision to make. Each “second mover” is given a $10 endowment. “First movers” are given endowments in amounts equal to the amounts kept (i.e. *not* sent) by the first movers in treatment A. Furthermore, the “second movers” in treatment C are given additional dollar amounts equal to the amounts received by second movers in treatment A from the tripled amounts sent by the first movers in treatment A. The subjects are informed with a table of the exact inverse relation between the number of additional dollars received by a “second mover” and the endowment of the anonymously-paired “first mover.”

5. Experimental Procedures

The experiments are designed with three games in order to discriminate between the observable implications of other-regarding preferences and trust or reciprocity. Two experiments are reported that have different social contexts. The experiment with a "strong social context" involves a second decision task that follows the first decision task in a treatment involving game A, B, or C. The presence of the second task does not introduce a repeated game because subject anonymity and random matching make it impossible for any subject to acquire a reputation. The experiment with a "weak social context" does not involve a second task.

5.1 Strong Social Context

This experiment includes treatments with Games A, B, and C explained in section 4. The instructions for each treatment announce the existence of a second task but do not explain that it is a group decision task involving the investment game.[2] The experiment sessions are run manually (i.e., not with computers). At the end of a session, a coin is flipped in the presence of the subjects to determine whether task 1 or task 2 has monetary payoff. The payoff procedure is double blind: (a) subject responses are identified only by letters that are private information of the subjects; and (b) monetary payoffs are collected in private from sealed envelopes contained in lettered mailboxes. A detailed explanation of the experiment procedures is contained in appendix 2. An additional appendix, available on request to the author, contains the instructions given to the subjects, the forms used during the experiment, and the questionnaires.

The experiment involves three treatments. Treatment A_{SC} implements game A, the investment game of BDMc. Each individual in the second-mover group is given 10 one-dollar certificates as a show-up fee. Each individual in the first-mover group is given 10 one-dollar certificates as a show-up fee and given the task of deciding whether he/she wants to transfer to a paired individual in the other group none, some, or all of his/her show-up fee. Any amounts transferred are tripled by the experimenters. Then each individual in the second-mover group is given the task of deciding whether she/he wants to return some, all, or none of the tripled number of certificates received to the paired individual in the other group. Treatment B_{SC} implements game B, which involves a decision task that differs from Treatment A_{SC} only in that the individuals in the "second-mover" group do not have a decision to make; thus they do not have an opportunity to return any tokens that they receive. Treatment C_{SC} implements game C(n), which involves a decision task that differs from Treatment A_{SC} as follows. First, individuals in the "first-mover" group do not have a decision to make. Each individual in the second-mover group is given 10 one-dollar certificates as a show-up fee. Individuals in the "first-mover" group are given show-up fees in amounts equal to the amounts kept (i.e. *not* transferred) by the first-mover subjects in Treatment A_{SC}. The second-mover subjects in Treatment C_{SC} are given "additional certificates" in amounts equal to the amounts received by the second-mover subjects in Treatment A_{SC} from transfers by the first-mover subjects in the latter treatment. The subjects are informed with a table of the exact inverse relation between the number of additional certificates received by a second-mover subject and the show-up fee of the paired "first-mover" subject.

The double blind payoffs are implemented by having each subject draw a sealed envelope containing a lettered key from a box containing many envelopes. If the coin flip selects task 1 for monetary payoff, the individuals use their task 1 keys to open lettered mailboxes that contain their monetary payoffs (from task 1 decisions) in sealed envelopes. There is no interaction between experimenters and subjects during decision-making parts of an experiment session. All distribution and collection of boxes containing envelopes for certificates kept and transferred is done by an assistant that is randomly selected from the subject pool. The assistant also flips the coin and observes all actions of the experimenters in calculating subjects' monetary payoffs.

All of the above design features are common information given to the subjects except for one item. The subjects in Treatment C_{SC} are *not* informed that the amounts of the "first-mover" show-up fees and the second-mover additional certificates are determined by subject decisions in Treatment A_{SC}.

All of the experiment sessions end with each subject being paid an additional $5 for filling out a questionnaire. First-mover subjects and second-mover subjects have distinct questionnaires. The questions asked have three functions: (a) to provide additional data; (b) to provide a check for possible subject confusion about the decision tasks; and (c) to provide checks for possible recording errors by the experimenters and counting errors by the subjects. Subjects did *not* write their names on the questionnaires. The additional data provided by the questionnaires include the subjects' reports of their payoff key letters. Subjects are asked to explain the reasons for their decisions. Data error checks provided by the questionnaires come from asking the subjects to report the numbers of tokens transferred, received, and returned. These reports, together with two distinct records kept by the experimenters, provide accuracy checks on data recording.

5.2 Weak Social Context

This experiment includes Treatments A_{WC}, B_{WC}, and C_{WC} , the implementation of Games A, B and C in a weak social context. The weak social context is one in which there is no second task. These data were previously reported in Cox (forthcoming). Except for elimination of the second task, Treatments A_{WC}, B_{WC} and C_{WC} are the same, respectively, as Treatments B_{SC} and C_{SC}.

6. Discriminating Between Altruism and Trust or Reciprocity

Treatment B differs from treatment A only in that the "second movers" do not have a decision to make; thus they do not have an opportunity to return any part of the tripled amounts sent to them. Since "second movers" cannot return anything in treatment B, first movers cannot be motivated by trust that they will do so. In contrast, in treatment A the first movers may be motivated to send positive amounts by both trust and altruistic other-regarding preferences. Thus conclusions about whether first-mover transfers in the investment game (treatment A) are partially motivated by trust are based on the difference between treatments A and B in the amounts of money sent by first movers to second movers.

Since "first movers" cannot send anything in treatment C, "second movers" cannot be motivated by positive reciprocity, that is, a need to repay a friendly action by a first mover. In contrast, in treatment A, second movers can be motivated to return positive amounts by reciprocity or by unconditional other-regarding preferences. Thus conclusions about whether second-mover transfers in the investment game are partially motivated by reciprocity are based on the difference between treatments A and C in the amounts of money returned by second movers to first movers.

As with any data, one needs a maintained theoretical model to interpret the data from the investment game triadic experiment. I begin by discussing the implications of a model of preferences over outcomes that can be conditional on the behavior of another. This model provides clear testable hypotheses about trust and reciprocity. Subsequently, I discuss some questions that have been raised about this approach.

Note that the definition of reciprocity in section 2 incorporates a possible dependence of preferences over outcomes upon the process that generated those outcomes and beliefs about the behavior of others. Such dependence can provide an explanation of why rational agents undertake actions involving trust and reciprocity. Thus, a first mover can rationally undertake a trusting action if she believes that this choice may trigger a social norm in the second mover that causes him not to defect. Alternatively, a first mover can rationally undertake a trusting action if he believes that the second mover has altruistic or inequality-averse unconditional other-regarding preferences. The experimental design for game triads explained in section 4 makes it possible to discriminate between the implications of unconditional other-regarding preferences and trust or reciprocity.

I will use the following specific criteria for deciding whether a first mover's behavior is trusting. A first mover will be said to undertake an action in the investment game that exhibits trust if the chosen action: (*i*) gives a positive amount of the first mover's money endowment to the second mover; and (*ii*) is risky for the first mover, in the sense that the amount of money that is sent is larger than the amount that would maximize the first mover's utility if none were to be returned by the second mover. Thus a trusting action requires a belief by the first mover that the second mover will not defect and keep too much of the profit generated by the first mover's decision to send a positive amount. If a first mover has self-regarding preferences then the act of sending any positive amount implies trust because such a first mover will lose utility if the second mover does not return at least as much money as the first mover gave up. But a first mover may have other-regarding preferences. Since, in the investment game any amount sent by the first mover is tripled, a first mover with altruistic preferences might prefer to give the second mover some money even if she knew that she would get nothing back. Thus the mere act of sending a positive amount of money is not evidence of trusting behavior unless it is known that first movers have self-regarding preferences. But the treatment B dictator game, together with the treatment A investment game, permit one to identify trusting actions, as follows.

Assume that each subject in every pair has preferences over her own and the paired subject's money payoffs that can be represented by a utility function. These preferences can be other-regarding or self-regarding. If the preferences are self-regarding then the utility function is a constant function of the other's money payoff. If the preferences are other-regarding then they can be altruistic or inequality-averse. In treatment B, a first mover chooses an amount to send from the set, S of integers weakly between 0 and 10. The choice in treatment B, s_b implies

(1) $$u^1(10 - s_b, 10 + 3s_b) \geq u^1(10 - s, 10 + 3s), \text{ for all } s \in S.$$

Now assume that the amount of money that the first mover gives to the second mover in treatment A, s_a is larger than the amount given in treatment B. Then we can conclude that the first mover has exhibited trust because the amount sent in treatment A is too large to be fully explained by other-regarding preferences. Thus, if $s_a > s_b$ then we know that the first mover is exposed to risk from the possibility that the second mover will defect and appropriate too much of the money transfer. Specifically, if the second mover were to return nothing in the event that

$s_a > s_b$, then statement (1) and strict quasi-concavity of u^1 imply that the first mover will have lower utility than he could have attained if he had known that the second mover would return nothing:

(2) $$u^1(10 - s_a, 10 + 3s_a) < u^1(10 - s_b, 10 + 3s_b)$$

because $s_a \in S$.

Next consider the question of identifying reciprocal behavior. The preferences over payoff (ordered) pairs can be conditioned on a social norm for reciprocity. For example, if the first mover in the investment game sends the second mover some of her money, the second mover may be motivated by a social norm for reciprocity to repay this generous action with a generous response. Within the context of a model of preferences over material payoffs, a social norm for reciprocity can be introduced with a state variable. Thus, the preferences over payoffs can be conditional on a state variable for reciprocity. This is an appropriate representation because, *if* there is reciprocal behavior, *then* individuals behave as if they are more altruistic towards another person after that person has been kind, generous, or trusting. The empirical question is whether or not second movers in the investment game choose more generous actions, after the first mover has intentionally sent them money, than they would in the absence of the first mover's action but the presence of the same money allocation.

When analyzing data from this experiment, I will use the following specific criteria for deciding whether a subjects' behavior is reciprocal. A second mover will be said to undertake an action that exhibits positive reciprocity if the chosen action: (*i*) returns to a generous first mover a positive amount of money; and (*ii*) is costly to the second mover, in the sense that the amount returned is larger than the amount that would maximize the second mover's utility in the absence of the generous action by the first mover. A second mover with self-regarding preferences will not return any money to the first mover. But a second mover with either altruistic or inequality-averse other-regarding preferences may return money to the first mover who, after making a positive transfer to the second mover, now has a lower money endowment than the second mover. Thus the mere fact that the second mover returns money to the first mover is not evidence of positive reciprocity. But the treatment C dictator game, together with the treatment A investment game, permit one to identify reciprocal actions, as follows.

A "second mover" in treatment C is given an endowment that is inversely related to the endowment of the paired subject. The endowments of a pair of subjects in treatment C are determined by a (distinct) first mover's decision in treatment A (but the subjects do not know this). Thus, the endowments of a pair of treatment C subjects are given by $(10 - s_a, 10 + 3s_a)$. In treatment C, a "second mover" chooses an amount to return from the set, $R(s_a)$ that contains the integers weakly between 0 and $3s_a$. The choice in treatment C, r_c implies

(3) $u^2(10 + 3s_a - r_c, 10 - s_a + r_c) \geq u^2(10 + 3s_a - r, 10 - s_a + r)$, for all $r \in R(s_a)$.

Suppose that the second mover returns to the first mover in the investment game a positive amount of money or, perhaps, even a larger amount than the first mover sent: $r_a \geq s_a$. This, in itself, does not support a conclusion that the second mover was motivated by positive reciprocity because the assumed choice could have been motivated by maximization of unconditional altruistic or inequality-averse other-regarding preferences. However, if one observes that $r_a > r_c$ then he can conclude that the second mover was motivated by reciprocity because the amount of money returned is too large to be fully accounted for by unconditional other-regarding preferences. This follows from noting that $r_a > r_c$, statement (3), and strict quasi-concavity of u^2 imply

(4) $u^2(10 + 3s_a - r_a, 10 - s_a + r_a) < u^2(10 + 3s_a - r_c, 10 - s_a + r_c)$

because $r_a \in R(s_a)$.

It might, at first, seem inconsistent with utility maximization for a subject to return an amount of money, r_a that satisfies inequality (4). But a social norm for reciprocity can change an agent's preferences over material payoffs. Such a norm can be incorporated into a theory of utility by introducing the possibility that an agent's preferences over outcomes can depend on the observed behavior of another. Specifically, with respect to reciprocity, an agent's preferences over his own and another person's material payoffs can depend on whether the other person intentionally helped him or intentionally hurt him or did neither. Thus, let λ_a be a state variable that depends on the amount of money sent by the first mover to the second mover in treatment A:

(5) $\lambda_a = f(s_a)$.

The utility to the second mover of the monetary payoffs in the investment game can be conditional on the reciprocity state variable. Thus there need be no inconsistency between inequality (4) and the norm-conditional-preference inequality,

(6) $u^2_{\lambda_a}(10+3s_a - r_a, 10 - s_a + r_a) \geq u^2_{\lambda_a}(10+3s_a - r, 10 - s_a + r)$, for all $r \in R(s_a)$.

Furthermore, experiments on reciprocal behavior can be characterized as research on the comparative properties of norm-unconditional (u^2), and norm-conditional ($u^2_{\lambda_2}$) utility-maximizing behavior.

A complete model for interpreting data from the triadic investment game experiment is presented in the appendix to Cox (forthcoming). Theoretical models that incorporate other-regarding preferences over outcomes that can be conditional on the perceived intentions of others are reported in Falk and Fischbacher (1999), Charness and Rabin (forthcoming), and Cox and Friedman (2002).

7. Comparison with the BDMc Data

One question of interest is whether the first- and second-mover transfers reported by BDMc are robust to the procedures used in the experiments reported here. Data from Treatment A_{SC} will be compared with data from the no-history (NH) and social-history (SH) treatments reported by BDMc.

7.1 Treatment A_{SC} vs. No-History BDMc Data

Treatment NH included 32 subjects run in three experimental sessions. Treatment A_{SC} included 30 subjects run in two sessions. Both data sets exhibit large variability across subjects. The amount sent varies from 0 to 10 and the amount returned varies from 0 to 20 in the data for both experiments. There are some differences between the two data sets. As reported in the first and second columns of Table 1, on average the subjects in Treatment A_{SC} both sent more ($6.00 vs. $5.16) and returned more ($7.17 vs. $4.66) than the subjects in Treatment NH. On average, the sending (or first-mover) subjects in Treatment A_{SC} made a $1.17 profit and those in Treatment NH made a $0.50 loss. Also, Treatment NH data are noisier than Treatment A_{SC} data in that the former have higher standard deviations for amounts sent and returned and the difference between amounts returned and sent.

The next to last row of Table 1 reports two-sample *t*-tests for differences of means and (non-parametric) Smirnov tests comparing Treatment A_{SC} and Treatment NH data. The .840 difference between mean amounts sent is not significantly different from 0 ($p = .234$) according to the *t*-test. The maximum difference between the cumulative distributions of amounts sent in Treatment A_{SC} and Treatment NH is not significant ($p > .864$) according to the Smirnov test. Since amounts returned are dependent on amounts sent, the next tests are on differences between amounts returned and amounts sent. The 1.67 mean difference between the Treatment A_{SC} and Treatment NH return minus sent amounts is not significant ($p = .134$). Also, the Smirnov test does not find a significant difference between the cumulative distributions of return minus sent amounts in the two treatments ($p > .220$).

The first and second rows of Table 1 report tobit estimates of the relation between amounts returned and amounts sent in Treatment A_{SC} and Treatment NH. The estimated model is given by

(7) $$R_t = \alpha + \beta S_t + \varepsilon_t,$$

where R_t is the amount returned by the second mover in subject pair t and S_t is the amount sent by the first mover in pair t. The bounds for the tobit estimation are the bounds imposed by the experiment design:

(8) $$R_t \in [0, 3S_t].$$

One would expect that the cone created by these bounds might produce heteroskedastic errors. In order to allow for the possibility of heteroskedastic errors, the tobit estimation procedure incorporates estimation of the θ parameter in the following model of multiplicative heteroskedasticity:

(9) $$\sigma_t = \sigma e^{\theta S_t}.$$

The .055 estimate of the intercept for the Treatment A_{SC} data is not significantly different from 0 (p = .966). The 1.17 slope coefficient is significantly different from 0 (p = .000). Because its standard error is .275, the slope coefficient for the Treatment A_{SC} data is also significantly different from 2 at a 1% significance level. Thus, the subjects' return behavior is significantly different from the prediction of completely-selfish subgame perfect equilibrium ($\beta = 0$) and the equal-split fairness focal point ($\beta = 2$). Sending subjects did, on average, earn a profit on the amounts they sent, but $\hat{\beta}$ is not significantly different from 1, which is the prediction of the zero-loss fairness focal point. The .152 estimate of the parameter, θ of the heteroskedasticity model is significant (p = .004). The right-most column of Table 1 reports the results from a likelihood ratio test for significance of the fitted model. It is highly significant (p < .005) for the Treatment A_{SC} data.

The .667 estimate of the intercept for Treatment NH data is not significantly different from 0 (p = .692). The .687 slope coefficient for Treatment NH data is significantly different from 0 (p = .082). Because its standard error is .394, the slope coefficient is also significantly different from 2 at a 1% significance level. Thus, the subjects' return behavior is significantly different from the prediction of completely-selfish subgame perfect equilibrium ($\beta = 0$) and the equal-split fairness focal point ($\beta = 2$). Sending subjects did, on average, make a loss on the amounts they sent, but $\hat{\beta}$ is not significantly different from 1. The .140 estimate of the parameter of the heteroskedasticity model is not significant (p = .175). The result of the likelihood ratio test is insignificance ($p > .100$).

The next to last row of Table 1 reports tobit estimates of the model,

(10) $$R_t = \alpha + \beta S_t + \gamma D_t S_t + \varepsilon_t$$

where:

(11) D_t = 1 forTreatment A_{SC} data

= 0 for Treatment NH data.

This estimation uses the bounds and heteroskedasticticity model given by equations (8) and (9). The estimate of γ is not significantly different from 0 (p = .179), which provides additional support for the conclusion that the differences between the data from Treatment NH and Treatment A_{SC} are not statistically significant.

7.2 Treatment A_{SC} vs. Social-History BDMc Data

Treatment SH included 28 pairs of subjects run in three experimental sessions. Treatment SH data looks even more similar to Treatment A_{SC} data than does Treatment NH data, as might be expected because the former two treatments both involve stronger social contexts than does Treatment NH. No significant differences between the Treatment A_{SC} and Treatment SH data are detected by the various tests reported in Table 1.

7.3 Conclusions Supported by Data from All Three Treatments

Treatment A_{SC} data, Treatment NH data, and Treatment SH data all support these three conclusions: (a) the data are inconsistent with the completely-selfish subgame perfect equilibrium of zero transfers; (b) the data are inconsistent with the equal-split fairness focal point; and (c) the data are *not* significantly different from the zero-loss fairness focal point.

8. Trust and Other-Regarding Preferences

Data from Treatments A_{SC} and B_{SC} can be used to that for the significance of trusting behavior in the strong social context. The tests reported in Table 2 use data for all except four of the subjects in sessions with Treatment B_{SC}. Tests with other subject samples, and the reasons for deleting data for four subjects in the tests reported in Table 2, are presented in appendix 1.

The first row of Table 2 reports the mean amounts sent by first movers in Treatments A_{SC} and B_{SC}. The mean amount sent was slightly larger in Treatment A_{SC} than in Treatment B_{SC} but the difference is insignificant ($p = .389$). The result from the Smirnov test is similar to that from the *t*-test, with no significant difference between Treatments A_{SC} and B_{SC} ($p > .10$). Thus there is no support in the strong social context data that the first-mover subjects sent part of their show-up fees to the paired second-mover subjects *because of* a trust that the second movers would not defect.

The second row of Table 2 reports estimates of amounts sent due to trust and other-regarding preferences in a weak social context. These estimates are based on data previously reported in Cox (forthcoming). The Treatment A_{WC} mean amount sent is and the Treatment B_{WC} mean amount sent is \$3.63. The one-tailed *t*-test for difference in means is significant ($p =$). The Smirnov test indicates a significant difference at the level of confidence. Thus the weak social context data support the conclusion that the first mover's behavior was motivated by trust as well as altruistic other-regarding preferences.

9. Reciprocity and Other-Regarding Preferences

Data from implementations of Games A and C can be used to test for the significance of positive reciprocity as a motive for second mover's behavior.

The first row of Table 3 reports the mean differences for all subjects between amounts returned and amounts "sent" in Treatments A_{SC} and C_{SC}. On average, the amount returned in Treatment A_{SC} exceeded the amount sent by \$1.17. Thus the second movers did share the profit provided by the first movers' transfers. The average amount returned in Treatment C_{SC} was \$1.23 less than the amount "sent." The difference between the Treatment A_{SC} and Treatment C_{SC} outcomes is in the direction implied by positive reciprocity. The one-tailed *t*-test for difference in means implies that the difference between the return minus sent means for Treatments A_{SC} and C_{SC} is significantly greater than 0 (p = .017). In contrast, the Smirnov test does not find a significant difference between the cumulative distributions of return minus sent amounts in Treatments A_{SC} and C_{SC}.

The sixth through tenth columns of the first row of Table 3 report tobit estimates using data for all subjects of the parameters of the following relation between amounts sent and amounts returned in Treatments A_{SC} and C_{SC}:

(12) $$R_t = \alpha + \beta D_t S_t + \gamma S_t + \varepsilon_t ,$$

where

(13) D_t = 1 for Treatment A_{SC} data

= 0 for Treatment C_{SC} data.

This estimation uses the bounds and heteroskedasticity model given by equations (8) and (9). Note that $\hat{\beta}$ is the estimate of the effect of reciprocity on amount returned by second movers to first movers. We observe that $\hat{\beta}$ is positive and significantly greater than 0 (p = .002). Thus the data provide support for behavior involving positive reciprocity in a strong social context.

The second row of Table 3 reports estimates of amounts returned due to reciprocity and other-regarding preferences in a weak social context. These data were previously reported in Cox (forthcoming). On average, the amount returned in Treatment A_{WC} was *greater than* the amount sent by . Thus the second movers did not share the profit provided by the first movers' transfers. The average amount returned in Treatment C_{WC} was less than the amount "sent." The difference between the Treatment A_{WC} and Treatment C_{WC} outcomes is in the direction implied by positive reciprocity. A two sample *t*-test for difference in means implies that the difference between the return minus sent means for Treatments A_{WC} and C_{WC} is significantly greater than 0 ($p =$). Furthermore, the Smirnov test finds a significant difference between the cumulative distributions of return minus sent amounts in Treatments A_{WC} and C_{WC} ($p >$).

The second row of Table 3 also reports tobit estimates of the parameters of the model given by statements (8), (9), (12), and

(14) D_t = 1 for Treatment A_{WC} data

= 0 for Treatment C_{WC} data.

The estimate of β is *not* significantly greater than 0 ($p = .566$). Hence the data do not provide support for behavior involving positive reciprocity in a weak social context.

Berg, Dickhaut, and McCabe (1995) noted that first movers who send all $10, and perhaps also those who send $5, might elicit larger returns by second movers. This possibility was examined by introducing dummy variables for $5 and $10 amounts sent into the above tobit models. Neither of the estimated coefficients for the $5 and $10 dummy variables is significantly different from zero for either strong or weak social context data.

10. Implications for New Utility Theories

Several recent papers have made important extensions of utility theory motivated by the large body of experimental data that has been interpreted as providing evidence of the importance of relative monetary payoff and perceptions of others' intentions as significant determinants of subjects' responses. Authors of these papers have demonstrated their consistency with data from several types of experiments. But the triadic design in the present paper generates data that pose new challenges and opportunities for extending theory. In order to demonstrate the behavioral patterns that are not yet being modeled, I will now explore the consistency of four of the new models with data from experiments with the triadic design.

Fehr and Schmidt (1999) present a utility theory of "self-centered inequality aversion" that they demonstrate is consistent with a large body of data. The two-agent version of their model is a utility function of the form

(14) $$u_i(x_i, x_j) = x_i - \alpha_i \max(x_j - x_i, 0) - \beta_i \max(x_i - x_j, 0),$$

where $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$. This model is referred to as a model of *self-centered* inequality aversion because of the restrictions on α_i and β_i. In my Treatment B, the first movers choose the payoffs given by the arguments in the utility function in statement (1). Substitution of these arguments into the utility function in equation (14) gives us

(15) $$u_1(10 - s^b, 10 + 3s^b) = 10 - s^b - \alpha_i \max(4s^b, 0) - \beta_i \max(-4s^b, 0).$$

Clearly, (15) is maximized by s^b = 0. This is inconsistent with the behavior of 38 out of 38 subjects in Treatment B_{sc} and with 19 out of 30 subjects in Treatment B_{wc}.

It is not just the linear form of (15) that makes this model inconsistent with Treatment B data. Any nonlinear generalization of (15) that retained its defining characteristic of inequality aversion model would also be inconsistent with the data because any positive value of s^b provides the second mover with a larger monetary payoff than the first mover who chooses the value of s^b.

Bolton and Ockenfels (2000) develop a utility theory of "equity, reciprocity, and competition" intended to be consistent both with subjects' other-regarding behavior in experiments such as dictator games and their competitive behavior in market experiments. They demonstrate that their model of preferences has implications that are consistent with a large body of data. Is their model consistent with data from the triadic experiments? It is straightforward to see that the answer is "no," as follows. The Bolton and Ockenfels (2000, p. 171) motivation function has two arguments, own monetary payoff, y_i and relative monetary payoff, λ_i; that is

(16) $$v_i = v(y_i, \lambda_i).$$

In my Treatment B, these variables are functions of the amount sent by the first mover, s^b:

(17) $$y_1 = 10 - s^b;$$

(18) $$\lambda_1 = \frac{10 - s^b}{20 + 2s^b}.$$

Assumptions 2 and 3 in Bolton and Ockenfels (2000, pgs. 171-172) imply that

(19) $$v(10,1/2) > v(10 - s^b, (10 - s^b)/(20 + 2s^b)), \ \forall s^b > 0.$$

The implication that (19) has for choice of s^b is inconsistent with the choices 38 out of 38 subjects in Treatment B_{sc} and with the choices of 19 out of 30 subjects in Treatment B_{wc}.

The reason for the inconsistency between the Bolton and Ockenfels model and data from the triadic experiments is similar to that for a nonlinear generalization of the Fehr and Schmidt model. The reason is that the model implies that an agent will never give a positive amount of money to another whose resulting payoff will be higher than his own regardless of how low is the cost of initiating the transfer. But when the cost of giving a paired subject an increase of \$1 is only \$0.33, as in Treatment B, a large majority of subjects give positive amounts even though this makes the recipients' monetary payoffs higher than the donors' monetary payoffs.

Levine (1998) develops a theory of altruism in which players' utilities are linear in their own monetary payoff, y_i and in others' monetary payoffs, y_j, $j \neq i$. He demonstrates that his model is consistent with data from some ultimatum game and market experiments. For two player games, Levine uses linear (adjusted) utilities of the form

(20) $$v_i(y_i, y_j) = y_i + \frac{a_i + \lambda a_j}{1+\lambda} y_j,$$

where $a_k \in (-1,1)$ is agent *k*'s "coefficient of altruism" and $\lambda \in [0,1]$ is another parameter. Let agent *i* be Mover 1 in game B and agent *j* be Mover 2. Then substitution of the game B monetary payoffs into (20) yields

(21) $$v_1(10 - s^b, 10 + 3s^b) = 10 - s^b + \frac{a_1 + \lambda a_2}{1+\lambda}(10 + 3s^b).$$

The linearity of v_1 in s^b implies that the chosen amount will be either 0 or 10. This is inconsistent with the choices of 33 out of 38 subjects in Treatment B_{SC} and with 15 out of 30 subjects in Treatment B_{WC}. One can also show that the utility of Mover 2 in game C, $v_2(10 + 3s^a - r^c, 10 - s^a + r^c)$ is linear in r^c which implies that the chosen amount is either 0 or $3s^a$. This is inconsistent with the choices of 24 out of 30 subjects in Treatment C_{SC} and with 21 out of 30 subjects in Treatment C_{WC}. But the inconsistency between Levine's model and the data comes from the linear form of the model; it is not a fundamental inconsistency.

Rabin (1993) develops a theory of fairness equilibria based on the following representation of agents' utilities. Define a_i, b_j, and c_i, respectively, as the strategy chosen by player *i*, the belief of player *i* about the strategy chosen by player *j*, and the belief by player *i* about the belief by player *j* about the strategy chosen by player *i*. Rabin (1993, pp. 1286-7) gives the expected utility function for player *i* as

(22) $$U_i(a_i, b_j, c_i) = \pi_i(a_i, b_j) + \widetilde{f}_j(b_j, c_i) \cdot [1 + f_i(a_i, b_j)],$$

where $\pi_i(a_i, b_j)$ is the monetary payoff to player *i*, $\widetilde{f}_j(b_j, c_i)$ is players *i's* belief about how kind player *j* is being to him, and $f_i(a_i, b_j)$ is how kind player *i* is being to player *j*.

Rabin's model is intended for normal form games but it may nevertheless be of interest to ask whether a fairness model with these player utilities is consistent with extensive form game data. Mover 1's monetary payoff function for game B is

(23) $\pi_i(a_i, b_j) = 10 - s^b$.

Mover 1's kindness function for game B is

(24) $f_i(a_i, b_j) = \dfrac{3s^b - 15}{30}$.

Thus Mover 1' expected utility function for game B is

(25) $U_1(s^b, b_j, c_i) = 10 - s^b + \tilde{f}_j(b_j, c_i) \cdot [1 + \dfrac{3s^b - 15}{30}]$.

The linearity of U_1 in s^b implies that the chosen amount will be either 0 or 10. As noted above, this is inconsistent with the choices of most subjects. Similarly, this linear model is inconsistent with the choices of most subjects in game C. But the inconsistency between Rabin's model and the data comes from the linear form of the model; it is not a fundamental inconsistency.

11. Concluding Remarks

This paper reports experiments with game triads that include the investment game. Several researchers had previously established the replicable result that the majority of first movers send positive amounts and the majority of second movers return positive amounts in investment game experiments. This pattern of results, and results from many other non-market fairness experiments, are inconsistent with the special case of game theory in which players' utilities are increasing functions of only their own monetary payoffs in an experiment. This leaves the profession with the task of constructing a less restrictive theory of utility that can maintain consistency with the empirical evidence. But this task cannot be undertaken successfully unless we can discriminate among the observable implications of alternative possible motivations. The game triad experiments reported here make it possible to discriminate among the observable implications for subjects' choices of trust, reciprocity, and other-regarding preferences in the investment game.

The game triads are implemented in two different contexts referred to as the strong and weak social contexts. Use of the strong social context makes it possible to conduct the experiment in an environment that is a little less artificial than the usual single-decision laboratory environment and a little more like the environment of ongoing social interaction that exists outside the laboratory. In the strong social context, the subjects are informed that there will be a second decision task after completion of the implementation of one of the three games. They are further informed that one or the other task will be randomly selected for monetary payoff. Because of anonymity and random pairing it is not possible to establish a reputation, hence the existence of the second task does not create a repeated game. The second task does create a richer social context than would otherwise exist in a laboratory experiment. The idea of play of a one-shot game in a social context is illustrated by the naturally-occurring interaction between drivers on big-city roads. The experiment environment without the second task is referred to as the weak social context.

There has been some previous experimentation aimed at investigating the implications of varying the social context of an experiment. In the first investment game experiment, Berg, Dickhaut, and McCabe (1995) implemented both a no-history treatment and a social history treatment in which the subjects were given the data from a preceding no-history treatment. The social history treatment is a stronger social context than the no history treatment. Data from my investment game treatment in the strong social context (Treatment A_{SC}) appears to be more similar to data from their social history (SH) treatment than to data from their no history (NH) treatment. But the various tests reported in Table 1 do not detect any statistically significant differences between Treatment SH and Treatment A_{SC} data nor between Treatment NH and Treatment A_{SC} data.

The present experiments reveal some effects from varying the social context: the amounts sent and returned are higher in all three treatments in the strong social context than they are in the weak social context. In addition, the change from strong to weak social context has effects on the measured incidence of trusting and reciprocating behavior in the game triad. The experiment in the strong social context does not provide any evidence of trusting behavior. The strong context does provide evidence of reciprocity. In contrast, the experiment in the weak social context provides evidence of both positive reciprocity and trust.

Comparison of data from the dictator treatments in the game triads with data from other dictator experiments provides additional insight into the properties of the other-regarding preferences revealed by the subjects' decisions. In the (DB1 and DB2) double blind dictator experiments reported by Hoffman, McCabe, Shachat, and Smith (1994), the average amount sent to the paired subjects by the dictators was $1. In the B_{WC} dictator treatment reported in this paper, the average amount sent by the dictators was $3.60. The price to the dictator of buying an additional $1 of income for the paired subject was $1 in the Hoffman, *et al.* experiment and it is $0.33 in Treatment B_{WC} reported here. The implied (arc) price elasticity of demand for increasing the other subject's income is −1.13.

The nonzero price elasticity of demand in dictator games for (the commodity that is) another agent's income illustrates a reason why the new models that incorporate other-regarding preferences are inconsistent with data from the game triad experiments. The models of inequality aversion fail because they do not allow for the price elasticity of demand for altruism. Thus, they imply that one agent will never transfer money to another whose resulting monetary payoff will be higher than his own regardless of how low is the (own-income) price per dollar increase in the other's monetary payoff. The data call for a generalization that incorporates nonzero price elasticity of demand for altruism and relaxation of the self-centered feature of the models' other-regarding preferences.

Results from the triadic design for experiments with the investment game demonstrate the importance of conducting experiments with designs that can discriminate among the motivations that are alternatives to that of wanting only to maximize one's own money income. Play in this game is characterized by observations of trust and reciprocity and by robust observations of other-regarding preferences. This distinction is central to the implication that data from the investment game have for guiding the development of utility theory to increase the empirical validity of game theory. The implication is that the data demand the inclusion of other-regarding preferences in utilities that are conditional on social norms for reciprocity or beliefs about others' intentions. But this conclusion should not be misinterpreted as a claim that social norms and/or the perceived intentions of others are *generally* unimportant determinants of behavior. Intentions and norms appear to be significant determinants of behavior in some experimental environments but the few other studies that have used control treatments for intentions have found mixed results.

Sally Blount (1995) compared second mover rejections in a standard ultimatum game with second mover rejections in games in which the first move was selected randomly or by an outside party rather than by the subject that would receive the first mover's monetary payoff. She found lower rejection rates in the random treatment than in the standard ultimatum game but no significant difference between the rejection rates in the third party and standard games. Gary Charness ((1996) used Blount's control treatments in experiments with the gift exchange game. He found somewhat *higher* second mover contributions in the outside party and random treatments than in the standard gift exchange game. Bolton, Jordi Brandts, and Ockenfels (1998) experimented with an intentions-control treatment in the context of simple dilemma games. In the control treatment, the row player "chooses" between two identical rows of monetary payoffs. They found no significant differences between the column players' responses in the control treatment and the positive and negative reciprocity treatments.

Endnotes

- Financial support was provided by the Decision Risk and Management Science Program, National Science Foundation (grant number SES-9818561).

1. This is not an example of the "indirect reciprocity" modeled in M. Nowak and K. Sigmund (1998a,b). In their model of repeated play, agents are rewarded for having a reputation for generous behavior and penalized for having a reputation for ungenerous behavior. In contrast, in the driving example it is impossible to acquire a reputation.

2. The group data are reported in Cox (2002), as are data showing gender effects.

References

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, Reciprocity, and Social History." *Games and Economic Behavior*, July 1995, 10(1), pp. 122-42.

Blount, Sally, "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences." *Organizational Behavior and Human Decision Processes*, August 1995, 63(2), pp. 131-44.

Bolton, Gary E., Jordi Brandts, and Axel Ockenfels, "Measuring Motivations for the Reciprocal Responses Observed in Simple Dilemma Games." *Experimental Economics*, 1998, 1(3), pp. 207-19.

Bolton, Gary E. and Axel Ockenfels, "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review*, March 2000, 90(1), pp. 166-93.

Charness, Gary, "Attribution and Reciprocity in a Simulated Labor Market: An Experimental Study." Working Paper, University of California at Berkeley, 1996.

Cox, J. (2002). "Trust, Reciprocity, and Other-Regarding Preferences: Groups vs. Individuals and Males vs. Females," in R. Zwick and A. Rapoport, (eds.), *Advances in Experimental Business Research*, Kluwer Academic Publishers.

Cox, J. (forthcoming). "How to Identify Trust and Reciprocity," *Games and Economic Behavior.*

Cox, J., and Friedman, D. (2002). "A Tractable Model of Reciprocity and Fairness." Paper presented at the March 2002 meetings of the Public Choice Society and the Economic Science Association.

Cox, J., Sadiraj, K., and Sadiraj, V. (2002). "A Theory of Competition and Fairness for Egocentric Altruists," University of Arizona discussion paper.

Dufwenberg, Martin and Georg Kirchsteiger, "A Theory of Sequential Reciprocity." discussion paper, CentER for Economics Research, Tilburg University, 1999.

Falk, Armin and Urs Fischbacher, "A Theory of Reciprocity." Working Paper No. 6, Institute for Empirical Research in Economics, University of Zurich, 1999.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, August 1999, 114(3), pp. 817-68.

Levine, David K., "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, July 1998, 1(3), pp. 593-622.

Nowak, M.A. and K. Sigmund, "Evolution of Indirect Reciprocity by Image Scoring." *Nature*, 393, pp. 573-7.

Nowak, M.A. and K. Sigmund, "The Dynamics of Indirect Reciprocity." *Journal of Theoretical Biology*, 194, pp. 561-74.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83, 1993, pp. 1281-1302.

Table 1. Comparison of Treatment A_{SC} and BDMc Treatments NH and SH

Data	*Sent Mean*	*Returned Mean*	Ret. □ Sent Mean	Sent Means Test	Ret. – Sent Means Test	Sent Smirnov Test	Ret. – Sent Smirnov Test	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	LR Test
Tr. A_{SC}	6.00 [2.59]	7.17 [4.82]	1.17 [3.71]	…..	…..	…..	…..	.055 (.966)	1.17 (.000)	…..	.152 (.004)	25.1 (.000)
Tr. NH	5.16 [2.94]	4.66 [5.55]	-.50 [4.89]	…..	…..	…..	…..	.667 (.692)	.687 (.082)	…..	.140 (.175)	3.56 (> .100)
Tr. SH	5.36 [3.53]	6.46 [6.19]	1.11 [4.31]					.275 (.858)	1.11 (.001)	…..	.096 (.160)	27.1 (.000)
Tr. A_{SC} vs. Tr. NH	…..	…..	…..	.840 (.234)	1.67 (.134)	.574 (>.864)	1.01 (>.220)	.479 (.650)	.747 (.003)	.345 (.179)	.136 (.007)	23.0 (< .005)
Tr. A_{SC} vs. Tr. SH	…..	…..	…..	.795 (.430)	.057 (.955)	.833 (>.465)	.520 (>.923)	.163 (.861)	1.13 (.000)	.026 (.897)	.124 (.003)	53.6 (.000)

p-values in parentheses.
Standard deviations in brackets.

Table 2. Money Sent Due to Trust and Other-Regarding Preferences

Social Context	Treatment A_{SC} or NH Mean	Treatment B_{SC} or B_{WC} Mean	Means Test	Smirnov Test
Strong	6.00 [2.59] {30}	5.81 [2.73] {38}	.184 (.389) [1]	.100 (> .10) [1]
Weak	5.16 [2.94] {32}	3.63 [3.85] {30}	-1.76 (.042) [1]	1.34 (>.052) [1] (<.068) [1]

p-values in parentheses.
[1] denotes a one-tailed test.
Standard deviations in brackets.
Number of subjects in braces.

Table 3. Money Returned Due to Reciprocity and Other-Regarding Preferences

Social Context	Treatment A_{SC} or NH Ret. – Sent Mean	Treatment C_{SC} or C_{WC} Ret. - Sent Mean	Means Test	Smirnov Test	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\theta}$	LR Test
Strong	1.17 [3.71] {30}	-1.23 [4.77] {30}	2.07 (.017)[1]	.133 (>.10)[1]	-.002 (.998)	.427 (.002)[1]	.748 (.000)	.181 (.000)	23.0 (< .005)
Weak	-0.50 [4.89] {32}	-0.83 [5.15] {30}	-0.26 (.397)[1]	.451 (>.964)[1]	.665 (.565)	.167 (.566)[1]	.517 (.056)	.172 (.008)	27.2 (.000)

p-values in parentheses.
[1] denotes a one-tailed test.
Standard deviations in brackets.
Number of subjects in braces.

Appendix 1. Analysis of Data for Alternative Subject Samples

During the first session involving Treatment B_{SC}, the questionnaire responses of three subjects indicated that they did not fully understand the decision task. This caused the experimenters to revise the Treatment B_{SC} subject information form before running subsequent sessions with this treatment.

In the last session with Treatment B_{SC}, there was one repeat subject. An experimenter recognized the subject as a repeat subject shortly after the session began. The double blind experimental protocol was discreetly compromised in order to learn this subject's identification letter (but not her name). This was done by silently counting the number of questionnaires that were sequentially deposited by the subjects in a stack in a collection box. It was determined that this subject's questionnaire was in position *n* in the stack. The subjects were not aware this was being done and all of their decisions had already been completed at the time the counting was done. This was the last session in this treatment.

Table 2 uses data for all subjects except the three possibly-confused subjects and one repeat subject. An argument can be made that data for all subjects should be used. An argument can also be made that data for all of the subjects in the first session with Treatment B_{SC} (using the original subject information form) should be deleted. The first row of Table 2A reports analysis of data for all subjects. The second row reports results from analysis of the data that remains after excluding all subjects from the first session with Treatment B_{SC} and the repeat subject. The conclusions are the same as from the analysis reported in Table 2.

Table 2A. Money Sent Due to Trust and Other-Regarding Preferences

Subject Sample	Treatment A_{SC} Mean	Treatment B_{SC} Mean	Means Test	Smirnov Test
All	6.00 [2.59] {30}	6.14 [2.82] {42}	-.143 (.588)	.100 (> .10)[I]
Excl. First Session and Repeat	6.00 [2.59] {30}	6.23 [2.72] {26}	-.231 (.626)[I]	.167 (> .10)[I]

p-values in parentheses.
[I] denotes a one-tailed test.
Standard deviations in brackets.
Number of subjects in braces.

Appendix 2. Experiment Procedures

A.2.1. The Strong Social Context

The subjects first assembled in the sign-in room of the Economic Science Laboratory and recorded their names, student identification numbers, and signatures on a form. Then a monitor was chosen randomly from the subject sample (by drawing a ball from a bingo cage) and given the responsibility of ensuring that the experimenters followed the procedures contained in the subject instructions for calculating money payoffs. The monitor was paid $20 for this job. The other subjects were not informed of the amount of this payment in order to avoid the possible creation of a focal earnings figure. Next the subjects were randomly divided into two equal-size groups, Group X and Group Y and escorted into the large room of the Economic Science Laboratory. The procedures differed somewhat across the three treatments because of the properties of the experiment design. I will first explain in detail the procedures used in Treatment A_{SC} and, subsequently, explain how procedures differed in Treatments B_{SC} and C_{SC}.

Treatment A_{SC} involves the BDMc investment game modeled in subsection 4.1 above. In a Treatment A_{SC} session, the Group X subjects were seated at widely-separated computer terminals with privacy side and front partitions. (The computers were not used.) The Group Y subjects were standing at the back of the room at the beginning of the session with Treatment A_{SC}. Each Group Y subject was given an envelope labeled “my show-up fee” that contained ten Task One $1 certificates. Each subject and the monitor were given copies of the instructions for “Task One” (the individual decision task). Then an experimenter read aloud the instructions. The instructions for all treatments are contained in an appendix available upon request. After the reading of the instructions was completed, the Group Y subjects were escorted back to the sign-in room by one of the experimenters. (The Group X subjects had no further contact with the Group Y subjects until after all decisions in both decision tasks had been completed.) Then the Group X subjects were given the opportunity to raise their hands if they had questions. If a subject raised his hand, he was approached by an experimenter and given an opportunity to ask questions and receive answers in a low voice that could not be overheard by other subjects. When there were no more questions, the experimenter left the room and the monitor took over to conduct the first-mover individual decision task with the Group X subjects.

The monitor carried a large box that contained smaller boxes equal in number to the number of subjects. Each subject was given the opportunity to point to any remaining small box to indicate she wanted that one. (The boxes all looked the same to the experimenters.) A subject opened her box to find an envelope labeled "my show-up fee" that contained ten Task One $1 certificates. The box also contained an empty envelope labeled "certificates sent to a paired person in Group Y" and an envelope containing a lettered Task One mailbox key. Finally, the box contained a one-page form that summarized the nature of the first-mover individual decision task. This form and the corresponding forms for other treatments are contained in an appendix available upon request. All envelopes in the box were labeled with the letter on the mailbox key.

Subjects were given 10 minutes to complete this task. When a subject was finished, he put all of the envelopes except the key envelope back in the box and summoned the monitor to collect the box. The monitor then carried the large box full of small boxes into another room for data recording and the preparation of boxes for the Group Y, second-mover subjects. The monitor witnessed all data recording and Group Y box preparations.

While the boxes were being processed, one experimenter escorted the Group X subjects out a side door of the Economic Science Laboratory and down the hall to the breakout rooms of the Decision Behavior Laboratory. Next, another experimenter escorted the Group Y subjects into the Economic Science Laboratory to get ready for their second-mover decisions in the individual decision task.

The Group Y subjects were given boxes by the monitor. Each box contained an envelope with a lettered Task One mailbox key. The box contained two empty envelopes, one labeled "my certificates" and the other labeled "certificates returned to the paired person in Group X." The box contained the tripled number of certificates sent by the paired person in Group X and a form summarizing the decision task. The form is contained in an appendix available upon request. The Group Y subjects had to decide how many of the certificates to put in the envelopes labeled "my certificates" and "certificates returned to the paired person in Group X." The Group Y subjects were given 10 minutes to complete the task. When a subject was finished, she put all envelopes except the key envelope back in the box and summoned the monitor to collect it. The monitor then carried the large box of little boxes to another room and watched the data recording

The second-mover decisions in Task One were conducted simultaneously with the first-mover decisions in Task Two. The first-mover decisions in Task Two were made by three-person committees that were formed by the experimenter by the order in which the subjects entered the laboratory from the hallway. Thus, the first three subjects were assigned to be in the first committee, the next three in the second committee, and so on. Each committee was seated in its own small breakout room. Each member of each committee was given the written subject instructions for Task Two. Then an experimenter read aloud the instructions while all breakout room doors remained open. Subjects were then given the opportunity to indicate whether they had any questions. If there was a question, the experimenter entered the appropriate breakout room and closed the door before the question was asked and answered. When there were no more questions, the experimenter left and the monitor took over. The monitor permitted the members of each committee to point to a small box contained in a large box to indicate which remaining box the committee wanted. A committee's box contained an envelope labeled "our show-up fee" that contained 30 Task Two $1 certificates. The box also contained an envelope labeled "certificates sent to a paired committee in Group Y" and an envelope containing a lettered Task Two mailbox key. Finally, the box contained a one-page summary of the group decision task. The form is contained in an appendix available upon request. The committees were given 20 minutes to complete their tasks. When a committee was finished, it put all envelopes except the key envelope back in the box and summoned the monitor by opening the door to its breakout room. The monitor carried the large box full of little boxes to the processing room and watched the data recording and preparation of boxes for the Group Y committees. Next, an experimenter escorted the Group X subjects back to the sign-in room. After all of the Group X subjects were in the sign-in room and the door was closed, an experimenter escorted the Group Y subjects out a side door of the Economic Science Laboratory and down the hallway to the breakout rooms of the Decision Behavior Laboratory.

The Group Y subjects then made their Task Two, second-mover decisions. Each Group Y committee was given an envelope labeled "our show-up fee" that contained 30 Task Two $1 certificates. The procedures for reading instructions, answering questions, and the role of the monitor were like those for the first-mover, Group X subjects. Each Group Y committee's box contained the tripled number of certificates sent to it by the paired committee in Group X. The box also contained an envelope with a Task Two key, a summary instruction form, and two empty envelopes. The empty envelopes were labeled "our certificates" and "certificates returned to the paired committee in Group X."

After the Group Y committees finished their Task Two decisions, they were escorted back down the hall to rejoin the Group X subjects in the Economic Science Laboratory. Next, an experimenter flipped a coin in the presence of all of the subjects and the monitor. The monitor announced whether the coin came up heads or tails. If heads (tails) then each Task One (Two) $1 certificate was exchanged for one United States dollar. While the subjects' money payoffs were calculated, they filled out the questionnaires. In addition to the salient money payoff, each subject was paid $5 upon depositing her completed questionnaire in a box. After the questionnaires were completed the Group X subjects went together to obtain sealed envelopes containing their money payoffs from lettered mailboxes. They had been asked to exit the building after obtaining their envelopes and not to open their envelopes until out of the building. After the Group X subjects had left, the Group Y subjects obtained their payoff envelopes from the lettered mailboxes.

The procedures for Treatment B_{SC} differed as follows from the Treatment A_{SC} procedures explained above. The Group Y subjects did not make a decision in Task One. The procedures for Treatment C_{SC} differed as follows from those for Treatment A_{SC}. At the beginning of Task One, the Group Y subjects were seated in the Economic Science Laboratory and the Group X subjects were standing at the back. The Group X subjects did not make a decision in Task One.

A.2.2. The Weak Social Context

Treatments B_{WC} and C_{WC} were conducted in the same way, respectively, as Treatments B_{SC} and C_{SC} except there was no Task Two.