Economics Department

Economics Working Papers

The University of Auckland *Year* 2003

Trust and Trustworthiness in a Sequential Bargaining Game

Ananish Chaudhuri* Sarah Khan† Aishwarya Lakshmiratan‡
Anne-Laure Py** Lisa Shah††

*University of Auckland, a.chaudhuri@auckland.ac.nz
†Wellesley College,
‡Wellesley College,
**Wellesley College,
††Asia Society,

Trust and Trustworthiness in a Sequential Bargaining Game[1]

Ananish Chaudhuri[a, b, *]

Sarah Ali Khan[b]

Aishwarya Lakshmiratan[b]

Anne-Laure Py[b]

Lisa Shah[c]

[a] Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand, *ananishc@yahoo.com*

[b] Department of Economics, Wellesley College, 106 Central Street, Wellesley, MA 02481, U.S.A., *skhan@wellesley.edu*, *alakshmi@wellesley.edu*, *apy61@hotmail.com*

[c] The Asia Society, 725 Park Avenue, New York, NY 10021-5088, U.S.A., *lshah@asiasoc.org*

***** Corresponding author. **Postal Address**: Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand. Tel: 64-9-373-7599 (373-7999) extn: 7661. Fax: 64-9-373-7427 E-mail: ananishc@yahoo.com

[1] We would like to express our sincerest thanks to Sara Graziano for enormous assistance in running our experiments. We are very grateful to Craig Parks and Uri Gneezy for valuable feedback. We thank Paul Strand and seminar participants at Wellesley College for their comments and ideas. We also thank three anonymous referees of this journal and the Editor (J. Frank Yates) for providing very detailed feedback as a result of which the paper is substantially improved. We are grateful to Wellesley College for providing the funds for this study.

Trust and Trustworthiness in a Sequential Bargaining Game

Abstract

We use a two-person extensive form bargaining game to examine individuals' trusting and reciprocal behavior and how those relate to their scores on a trust survey. In keeping with prior research, we find that the "self-interested" outcome is rejected by a majority of individuals. People who score high on the trust survey are both trusting and are also trustworthy, in that they reciprocate others' trust. But, people with low trust scores often exhibit trust but are not trustworthy. These "inconsistent trusters" seem to be interested in exploiting the trust and trustworthiness of others in increasing their own payoff.

JEL Classification: *C72, C91*

PsycINFO classification*: 2343, 3020*

Keywords: Trust; Reciprocity; Social Values Orientation; Sequential Game; Bargaining

1. Introduction

The "homo-economicus" assumption of economics perceives humans as acting out of pure self-interest in economic transactions and considers deviation from self-interested profit maximization as irrational. Behavioral game theory however, has built a large precedent of empirical studies, which demonstrates that actual behavior often deviates from the self-interested predictions by exhibiting notions of trust and reciprocity.[2] In one such study looking at the role of trusting and reciprocal motivations, McCabe, Rassenti and Smith (1998) ask if the large amounts of trusting and reciprocal behavior they observe can be explained on the basis of pure trust, or if they require the prospect of punishment. The authors comment (1998, p. 22),

> "Reciprocity and its origins in trust and/or punishment is in need of being modeled to account for a variety of behavioral reputational types: those who offer cooperation on the basis of pure trust, those who require the prospect of punishment, those who defect when cooperation is offered and those who, faced with defection, tend to respond with punishment." [3]

In this paper we address this point raised by McCabe et al. (1998), examining if "pure trust" can be the basis of trusting behavior. We test to see if a trust score, calculated for each individual using the Social Values Orientation survey, used by Yamagishi (1986) and Yamagishi and Sato (1986), is correlated to trusting and reciprocal moves in a two-person extensive form game (described below) to explore the following questions. (1) Do trust and reciprocity based motivations affect the outcome in this game? (2) Can

[2] The literature in this area is voluminous. Representative publications are Berg et al. (1995), Chaudhuri et al. (2002), Cox (2001), Gneezy et al. (2000) and McCabe, et al. (1998, 2002). Camerer (2001, Chapter 3) provides a thorough review of much of the existing literature in economics. Kramer (1999) provides an extensive literature review from the perspective of psychology and organizational behavior.

[3] McCabe et al. (1998) is not the first paper to make this point. Many others have done so. Messick and McClintock (1968) is one of the earliest studies to make this point about different types of players. See also Kuhlman, Camac and Kunha (1986) for similar arguments.

trust be explained by what McCabe et al. (1998) refer to as "pure trust"? In other words are trust scores correlated to trusting moves? (3) Is the trusting move correlated to reciprocity? Do those who trust, also reciprocate the trust of others?[4] What we find is that by and large the "self-interested" outcome is rejected by a majority of individuals. People with high trust scores are both trusting and trustworthy in that they trust others and also reciprocate others' trust. But, people with low trust scores may show trust, but are often not trustworthy in that they often do not reciprocate the trust of others. Rather these "inconsistent trusters" seem interested in exploiting the trust and trustworthiness of others in increasing their own payoff.

In the game that we use (see Figure 1) individual rationality suggests one outcome but players may achieve a better monetary outcome if they are motivated by trust and reciprocity.[5] In this game, Player 1 moves first and chooses one of two strategies: Top or Bottom. Player 2 gets to see what Player 1 chose and responds by picking either Left or Right. The payoffs are denoted in dollars and cents and were paid to the subjects in cash at the end of the experiment. The subgame perfect outcome (using backward induction) in this game is for Player 1 to choose Top and Player 2 to choose Left. This nets $2.50 for Player 1 and $1.25 for Player 2.[6] This outcome is indicated in the figure as the SPE

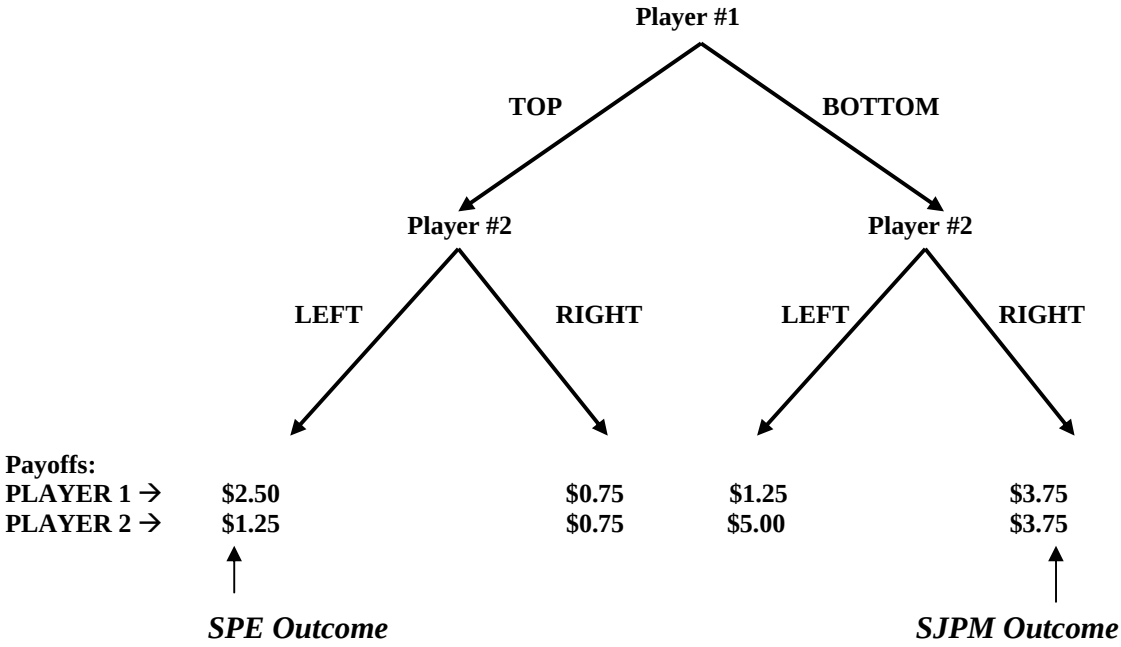
[4] Chaudhuri, Sopher and Strand (2002) fail to find any correlation between trusting and reciprocal moves, a finding echoed in Abbink et al. (2000).

[5] The game is really a simpler version of the ones used by McCabe et al. (1998, 2002).

[6] Suppose Player 1 chooses Top and places Player 2 at the left node. At this node a rational Player 2 will choose Left over Right since the former yields $1.25 while the latter yields $0.75. On the other hand if Player 1 chooses Bottom and puts Player 2 at the right node then Player 2 will choose Left again since that nets her $5.00 as opposed to Right which nets her $3.75. Thus knowing that Player 2 will respond with Left at either node, Player 1 should choose Top and put Player 2 at the left node, because in this case following Player 2's Left choice Player 1 gets $2.50. On the other hand if Player 1 chooses Right and puts Player 2 at the right node, a choice of Left by Player 2 gets Player 1 only $1.25. Player 1 makes more money in the {Top, Left} outcome making this is the subgame perfect outcome.

outcome. However there is a symmetric joint payoff-maximizing outcome at {Bottom Right} which nets $3.75 for each player. This is indicated as the SJPM outcome.

Figure 1: Payoffs of Player 1 and Player 2



The problem with the SJPM outcome is that in order to attain this Player 1 would have to make an explicitly trusting move of choosing Bottom at the beginning. The choice of Bottom requires trust because Player 2 can take advantage and respond with Left which gives $1.25 to Player 1 and $5.00 to Player 2. In that case Player 1 is better off at the SPE outcome where she gets $2.50. However if Player 2 reciprocates Player 1's trust by choosing Right after Bottom, then both players are better off, with each getting $3.75, more than what they would get if the SPE outcome is reached. That is Player 2 would have to think along the following lines. "I should choose Right in response to Player 1's choice of Bottom, and Player 1 is obviously trusting me to do so since she could easily have chosen Left and given me no more than $1.25 at the SPE outcome." There is one final outcome that is possible – the so-called "retributive" outcome – at

{Top, Right}. This is where Player 1 chooses Top (possibly in order to get to the SPE outcome) but player 2 looks at Player 1's choice and is dissatisfied in that Player 1 is obviously not trusting Player 2 enough to get to the SJPM outcome. Player 2 in turn decides to punish Player 1 (for her perceived lack of trust) by moving Right in response to Top which gets each player $0.75. Player 2 foregoes some payoff to punish Player 1 for her lack of trust – an act of "negative reciprocity" or "retribution". For a detailed discussion of this phenomenon see Abbink et al. (2000). McCabe, Rassenti and Smith (1998) and McCabe, Rigdon and Smith (2002) also address the issue of retribution but find little evidence of such behavior in their data.

Notice further (see Figure 2) that in this game, Player 2 has a dominant strategy, which is to play Left, making (Top, Left) the dominant strategy Nash equilibrium of the normal form game. According to Schotter, Weigelt and Wilson (1994), presenting the same game in extended form rather than the normal form makes the dominant strategy and the SPE outcome (or the Dominant Strategy Nash Equilibrium) more recognizable and increases the probability of reaching it. Schotter et al. (1994) found that the presentation or "framing" of the problem significantly affects the strategy chosen by individuals, and that using an extensive form results in more choices commensurate with the Nash Equilibrium. We thus present our subjects with an extended form game thereby giving the SPE outcome of (Top, Left) its best shot.

We proceed as follows. In the next section we describe the experimental procedure. After that we present our results in Section 3. Finally in Section 4 we offer some concluding thoughts and remarks.

Figure 2: The Game in Normal Form

	PLAYER #2		
PLAYER#1		LEFT	RIGHT
	TOP	2.50, ***1.25***	0.75, ***0.75***
	BOTTOM	1.25 , ***5.00***	3.75, ***3.75***

Left is a Dominant Strategy for Player 2

2. Experimental Procedure

76 subjects participated in the study. They were recruited from among Wellesley College students (who are all female) via postings on an electronic bulletin board. We took care to recruit only an even number of subjects for each session. They were equally separated into two rooms upon arrival and prior to the beginning of the game or the reading of instructions. After the subjects were situated into their rooms, each subject was given the instructions that indicated a unique subject identification number. This number was used to pair each subject with a distinct Player 1 or Player 2 partner in the other room. The subjects were first given the instructions necessary for filling out the trust survey. After everyone had filled out the trust survey, they were then given the instructions for playing the sequential game. The experiment lasted about 30 minutes and the average payoff was $6.20.

2.1 Trust Survey

Each subject was given a trust survey to understand the subject's "pure trust" levels and how this correlated with their Player 1 and Player 2 decisions. The survey used is the Social Values Orientation (SVO) Scale used by Yamagishi (1986) and Yamagishi and Sato (1986).[7] The SVO scale is considered a good predictor of one's predisposition to trust strangers and has been used by previous researchers to look at the correlation between trust and propensity to contribute to public goods[8] and that between trust and the propensity to cooperate in a prisoner's dilemma game.[9] The SVO scale is short and consists of five questions[10] which are each answered by choosing one of the five options: strongly disagree, disagree, neutral, agree and strongly agree. An answer of strongly disagree gets a 5 while strongly agree gets a 1 except for question 4 which is reverse scored. The lowest possible score is 5 (least trust) and the highest possible is 25 (most trust). The average score of the subjects in this study is 12.88 out of a possible 25 with a standard deviation of 3.286. The median score is 13 and the modal score is 13 as well with 15 people scoring 13 out of 25. The lowest score recorded for our study is 5 and the highest is 21.

2.2 Sequential Game Overview

In the sequential game each subject played both roles of Player 1 and Player 2, except each player interacts with a different player in each role. This preserves the one-

[7] This trust scale was originally developed by Yukawa (1985) based on the factor analysis of 60 items related to trust.
[8] See Yamagishi (1986) and Yamagishi and Sato (1986)
[9] See Parks et al. (1995, 1996)
[10] (1) Most people tell a lie when they can benefit by doing so. (2) Those devoted to unselfish causes are often exploited by others. (3) Some people do not cooperate because they pursue only their own short-term self-interest. Thus, things that can be done well if people cooperate often fail because of these people. (4) Most people are basically honest. (5) One should not trust others until one knows them well.

shot nature of the game allowing for no scope for reputation building. The following scheme explains the pairing process.

Room A Player 1	**Room B Player 2**	**Room B Player 1**	**Room A Player 2**
1	5	5	4
2	6	6	3
3	7	7	2
4	8	8	1

The game started with each subject making a Player 1 decision. Each subject recorded her move on the decision sheet and handed it in to the experimenter. The sheets were then passed to the designated partner in the other room. After all subjects recorded and looked at their partner's Player 1 decision, they responded with a Player 2 decision. This decision was recorded on the decision sheet and then passed back to their paired Player 1, so that the Player 1 would now know what her payoff was. Each subject was given a record sheet for recording her payoffs. A copy of the experimental instructions, along with the decision and record sheets, is provided in the appendix.

2.3 Hypotheses

Hypothesis 1: *Even though {Top, Left} is the subgame perfect outcome of the game, still we will observe a large number of {Bottom, Right} outcomes due to the presence of trusting and reciprocal motives.*

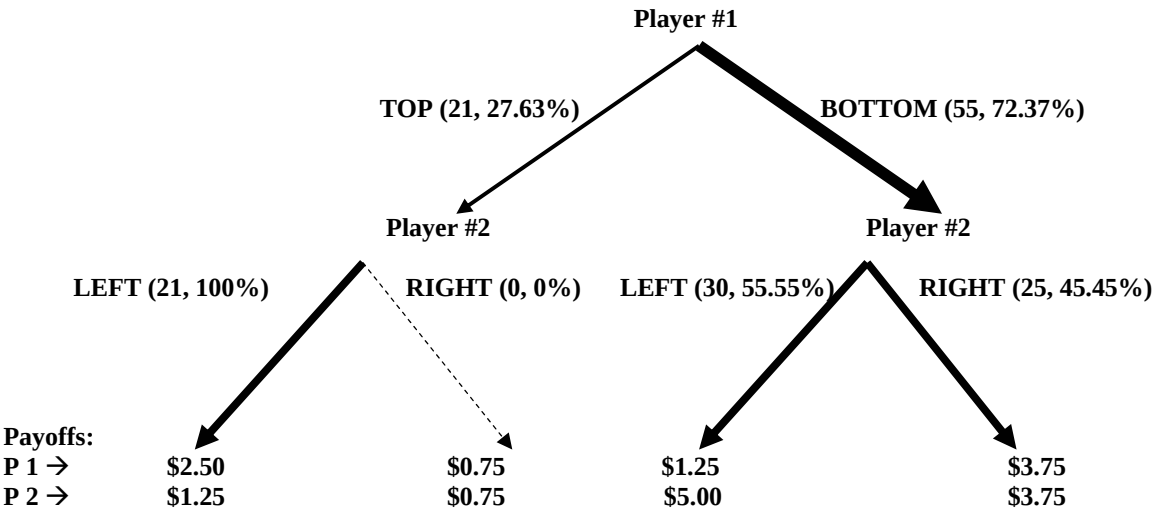
Hypothesis 2: *The trust score correlates with and can explain each subject's Player 1 and Player 2 decision because it explains their views towards trust and reciprocity.*

3. Results and Analysis

3.1 Overview

A summary of the results from the 76 plays of the game are presented in extended form below. The figures in parentheses are the number of players (out of 76) who chose this response and the percentage of players who chose this response, respectively. We have used thicker or thinner arrows to depict the relative frequencies of various moves.

Figure 3: Frequency of Moves



We define the different Player 1 choices and Player 2 responses according to the behavior they represent.

Player 1 choices: A Player 1 choice of Top was labeled as a Non-trusting move, while a Player 1 choice of Bottom was labeled as a Trusting move.

Player 2 choices: A Player 2 choice of Left in response to Top signaled reaching the Nash Equilibrium outcome, or the SPE outcome, which is the outcome predicted by rational choice and game theory. A Player 2 response of Right to Top suggests Player 2 is punishing Player 1 with lower payoffs for Player 1's lack of trust. This is labeled as a Retributive move. A Player 2 choice of Left in response to Bottom suggests non-

reciprocal behavior by Player 2 given that Player 1 played an initial trusting move. A Player 2 choice of Right in response to Bottom suggests reciprocal behavior with Player 2 reciprocating Player 1's initial trust attaining the Symmetric Joint Payoff Maximizing (SJPM) outcome.

As the results show, an overwhelming majority of players trusted in their first move (72.37%). The SPE outcome predicted by game theory is rejected in the first move by an overwhelming majority of the players. In keeping with the findings of McCabe et al. (1998, 2002), there is no evidence of retributive behavior at all with no subjects choosing Right as a response to Top. Reciprocal behavior however is less prevalent with about 45% (25 out of 55) of those who have the choice of reciprocating actually doing so with the other 55% (30 out of 55) choosing the non-reciprocal move of Left in response to Bottom.

3.2 Choice 1 decisions and Trust

Given that roughly 72% of the players chose the economically "irrational" move of Bottom as Player 1, we wanted to test if this choice was based on pure trust. To test this, we regressed Player 1 choices against individual trust scores on the survey using a binary probit regression. A Player 1 choice of Top was assigned a value of 0 and a choice of Bottom was assigned a value of 1. The coefficient of the trust score in the regression is –0.0133 (standard error = 0.074) and is not significantly different from zero. The corresponding test-statistic is 0.18 with a p-value of 0.857. It therefore appears that there is no statistical correlation between players' pure trust levels as measured by their survey responses, and their actual decision to choose the trusting move (Bottom) as Player 1 in the game.

3.3 Player 2 decisions, Trust and Reciprocity

Following a Player 1 choice of Top, Player 2 chose Left was in all 21 cases (100%) resulting in the SPE outcome. There was no retributive behavior seen (Right as a response to Top).

A more complicated picture is presented at Player 2's decision node following a choice of Bottom by Player 1. At this point, 55 players had to choose between Right (reciprocating Player 1's trust by settling for an equally high payoff for both) and Left (not reciprocating Player 1's trust and getting the highest payoff for oneself at the cost of lowering Player 1's payoff). The choice in this case was split almost half way with 25 out of 55 (45.45%) players reciprocating trust and reaching the SJPM outcome, and the other 30 out of 55 (55.55%) not reciprocating. To see what drove these responses, a probit regression was run on the 55 Player 2 responses to a Player 1 choice of Bottom to see if it was correlated with players' trust scores (higher trust score leading to greater reciprocal behavior). However we find that the trust score is not a significant predictor of the Player 2 responses. The coefficient of the trust score variable is 0.0633 (standard error = 0.0513). The corresponding test-statistic is 1.23 with a p-value of 0.217.

While the probit results are not significant, there is some evidence that people with higher trust scores are more likely to reciprocate a trusting move by their pair members. We find that among the 76 subjects there were 46 subjects who scored above the average of 12.88 while 30 subjects scored below the average. Let us label the former group "high trusters" and the latter "low trusters".[11] We find that out of the 46 "high

[11] The usual practice is to do a median split and look at those who are above the median and those who are below. However as discussed in Section 2.1 there are 15 subjects who scored the exact median of 13 on the trust survey. It would be hard to argue for their inclusion in either the group above or below the median and thus we would have to exclude 15 data points. Thus we have chosen to look at subjects above and below

trusters" 18 subjects chose the reciprocal move of Right when faced with the trusting move of Bottom. For the 30 "low trusters" only 7 responded with Right to Bottom. Using a sample proportions test[12] we get a z-value of 1.6 which is significant at the 0.06 level using a 1-tailed test but narrowly misses conventional levels of significance using a 2-tailed test.

However, as we discuss in the next section, there is a much more subtle phenomenon that is going on here which is not captured by simply looking at the relation between trust scores and reciprocity for Player 2. Given the results that we present in the next section it is not very surprising that we fail to get high correlation between trust scores and reciprocity. Let us turn to this next.

3.4 Consistency between Trust and Reciprocity

One very curious aspect of the data is the large inconsistency between 72% of players making the trusting move as Player 1 and only 45% of players reciprocating that trust as Player 2. Nearly half of the players therefore seem to have different behavioral allegiances in each play of the game. 30 out 55 subjects behave inconsistently in that they choose to trust in their first move but do not reciprocate others' trust in their second move. 25 out of 55 players, however, are consistent in choosing the trusting move as Player 1 and the reciprocal move as Player 2.

---

the mean instead. Later on in Section 2.5 where we focus on individual behavior we compare subjects on the basis of both the mean and the median.

[12] If the two relevant sample proportions are p_1 (18 out of 46 in this case) and p_2 (7 out of 30), and the two samples have n_1 (=46) and n_2 (=30) members respectively, then the corresponding test-statistic is

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 * (1 - p_1)}{n_1} + \frac{p_2 * (1 - p_2)}{n_2}}} = 1.6$$

We would like to know if there are significant differences in the trust levels of these two groups. In order to look at this, we divide the sub-sample of 55 players faced with a reciprocity choice into those who were consistent with their trust (who had trusted as Player 1 and reciprocated as Player 2) and those who were inconsistent (who had trusted as Player 1 but not reciprocated as Player 2). We then compare the trust scores of these two groups using both parametric and non-parametric tests. Table 1 presents the results. The consistent players have an average trust score of 14 while the inconsistent ones average 11.61. Using a t-test, we find a statistically significant difference (at the 0.02 level using a 1-tailed test and at the 0.03 level using a 2-tailed test) between the average trust scores of those who were consistent in their choices and of those who were inconsistent. The result is even stronger using the non-parametric Wilcoxon test, which returns a z-value of 2.97 with a corresponding p-value of less than 0.01. The inconsistent players do make a trusting move as Player 1 but are not trustworthy in the sense that they do not reciprocate a trusting move as Player 2. In the next section we discuss why, we think, these players behave in an inconsistent manner.

There is another way of looking at this distinction between consistent and inconsistent player. Let us look at the sub-sample of subjects who chose the trusting move of Bottom as Player 1. Then let us look at how many of those subjects faced a trusting move of Bottom as Player 2. This leaves us with 39 subjects. The question is does the trust score predict well how these players behaved as Player 2? The answer is an emphatic yes. We run a binary probit regression with Player 2's choice on the left hand side and the trust score on the right hand side for those players who chose to trust as Player 1 and had the option of reciprocating others' trust as Player 2. A choice of Left by

Player 1 is assigned a value of "0" while a choice of Right gets a value of "1". Table 2 presents the results. We find that a high trust score makes it much more likely that the player would be trustworthy. An increase in the trust score significantly increases a subject's probability of reciprocation. If we compute the marginal effects then we find that a 1-point increase in the trust score increases the probability of reciprocation by 5.4%. Thus subjects with high trust scores are more likely to choose the trusting move of Bottom as Player 1 and to reciprocate as Player 2.

Table 1: Trust scores of consistent and inconsistent players

	Trust as Player 1 and reciprocate as Player 2 **Consistent**	Trust as Player 1 but not reciprocate as Player 2 **Inconsistent**
Number of people	21	18
Average trust score	14	11.61
Variance	8.5	14.72
t-statistic	2.2067	
p-value (2-tailed)	0.03	
p-value (1-tailed)	0.02	
Wilcoxon z-statistic	2.97	
p-value	0.003	

Table 2: Probit Regression
Dependent Variable: Probability of Reciprocation

	Coefficient	Standard Error	test-statistic	p-value
Trust Score	0.137	0.066	2.08	0.037
Constant	-1.67	0.879	-1.90	0.057
Pseudo-R^2	0.09			

3.5 Analysis of Individual Behavior

Probing a little deeper into the choices of players with high and low trust scores we find the following. There are 46 players who scored more than the average of 12.88

(the "high trusters") on the trust survey and 30 who scored less than the average (the "low trusters"). Among the high trusters, there were 16 people (34.8%) who were consistent in that they chose the trusting move of Bottom as Player 1, and the reciprocal move of Right against Bottom as Player 2. Among the low trusters, only 5 out of 30 (16.7%) were consistent. If we use a sample proportions test, then this difference in consistency is significant at the 0.08 level using a 2-tailed test and at the 0.05 level using a 1-tailed test. The corresponding z statistic is 1.81, which corresponds to 46.49% of the area under the normal curve.[13]

As we mentioned above the usual practice is to do a median split and look at those above and below the median. This is problematic for us since there are 15 people who scored the median of 13 and we would have to exclude them in any median split analysis. If we do carry out a median split then we find the following. Excluding those 15 subjects who are at the median, we get 30 subjects who scored below the median and 31 who scored above the median. Out of the 30 subjects below the median 5 were consistent while out of the 31 above the median 11 were consistent. Once again using a sample proportions test we get a test-statistic of 1.71 which corresponds to 45.64% of the area under the normal curve and is thus significant at 0.09 level using a 2-tailed test and at the 0.05 level using a 1-tailed test.

We find that in the first move roughly the same proportion of "high trusters" and "low trusters" (we are using the mean score as the cut-off between high and low trust so

[13] If the two relevant sample proportions are p_1 (16 out of 46 in this case) and p_2 (5 out of 30), and the two samples have n_1 (=46) and n_2 (=30) members respectively, then the corresponding test-statistic is

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 * (1 - p_1)}{n_1} + \frac{p_2 * (1 - p_2)}{n_2}}} = 1.81$$

that we can include all 76 subjects) chose the trusting response of Bottom. 22 out of 30 "low trusters" (73.33%) and 33 out of 46 "high trusters" (71.7%) chose Bottom as Player 1. In the second move however, we see more reciprocal behavior from the high trusters. As we saw in the previous section, players with consistent choices in both moves had significantly higher pure trust levels than players who were inconsistent in their choices between moves. Those who reciprocated as Player 2 therefore were more likely to have been driven by their pure trust levels compared to those who trusted at the first move but then maximized their payoff at the second move with no consideration of reciprocity. These low trusters seem less motivated by pure trust but rather are interested in exploiting the trust and trustworthiness of others in increasing their own payoff. It seems to us that the low trusters engage in the following course of action. As Player 1 they repose trust on the other player hoping for reciprocity from her and consequently the bigger (SJPM) payoff of $3.75. However as Player 2 (and facing a trusting move of Bottom from Player 1) they choose not to reciprocate and choose Left, thereby grabbing the larger payoff of $5.00 for themselves. This yields these players the maximum possible payoff of $8.75. In the event of arriving at the SJPM outcome twice the maximum a player can get is $7.50 ($3.75 times 2). Thus by taking a calculated risk these low trusters can increase their payoff by $1.25.

Thus we find that pure trust levels do not explain all the Player 1 choices, some of which are motivated by taking advantage of the trust and trustworthiness of others. Player 2's decision to reciprocate, in and of itself, is not explained very well by pure trust levels either. However, consistency in trust and reciprocity seems to be correlated with pure

trust levels. Players who have high trust scores tend to be more consistent in trusting as Player 1 and reciprocating as Player 2 than players with low trust scores.

We conclude this section by making the following observation. Based on our findings we can say that the Yamagishi SVO scale (and possibly other standard surveys of this type such as the General Social Survey) can be used as a good predictor of trustworthy behavior but not necessarily trusting behavior. Glaeser et al. (2000, p. 841) also make this point when they say "standard survey questions about trust do not appear to measure trust. However they do measure trustworthiness, which is one ingredient of social capital."

3.6 Discussion of our results

In the current study we have identified two different groups of subjects. One group (characterized by high trust scores) offers trust and reciprocity unconditionally. The other group (characterized by low trust scores) is more self-interested and seems to be looking for ways to exploit the trust and reciprocity of others to enhance monetary payoffs. The latter group's decision to make a trusting move initially by choosing Bottom is in the nature of a gamble or calculated risk and is based on calculations of expected reciprocation by the pair member thereby maximizing monetary gains. These decisions then have more to do with calculated risk-taking rather than trust per se. See Eckel and Wilson (2002) and Chaudhuri and Gangadharan (2002) for more discussions along these lines. Both these papers find that for some people the decision to trust is related to risk attitudes. Thus one group in our subject pool operates on the basis of "pure trust", people who are both trusting and trustworthy. The second group consists of those who display trust more as a calculated gamble and do not reciprocate the trust reposed in them by their

pair members. The trust exhibited by the former group "can be conceptualized as a social orientation toward other people and toward society as a whole" while for the latter group their trust can be "conceptualized more as a calculative orientation towards risk." (Kramer, 1999, p. 3). This latter group of subjects, to use McCabe et al.'s (1998) idea, would probably require the existence of an explicit punishment mechanism in order to reciprocate the trust of others. They would reciprocate only if the person who trusted them in the first place can punish them in case of non-reciprocity.

The nature of the beliefs that players hold about their pair members is also crucially important. Geanakoplos, Pearce and Stacchetti (1989) introduce the paradigm of psychological games where the payoff to each player depends not only on what every player does but also on what she thinks every player believes, and on what she thinks they believe others believe and so on. By explicitly introducing players' beliefs, the authors argue that in many cases the psychological payoffs associated with a terminal node are endogenous and so are the equilibrium strategies. Geanakoplos et al. (1989, p. 61) suggest that along the same lines players' utilities may depend "not only on the physical outcome of the game but also on their beliefs before and during play…. Consequently, the traditional theory of games is not well suited to the analysis of such belief dependent psychological considerations as surprise, confidence, gratitude, disappointment, embarrassment and so on." They go on to say (on p. 63) that "the failure of backward induction in psychological games results from the fact that when a node is reached, it does not capture adequately the state of the game: the node identifies a history of play, but not the players' beliefs."

What this implies in terms of the present study is that if we allow for different distributions of beliefs on the part of players, {Top, Left} is not the only possible equilibrium outcome. Consider the following very rudimentary example.

Suppose each player 1 believes that with probability "p" the player 2 she is paired with is trustworthy while with probability "1-p" she is not. A trustworthy player 2 always chooses Right in response to Bottom while an untrustworthy player 2 always chooses Left. Then it is easy to see that the expected payoff to Player 1 from choosing Bottom is $1.25(1-p) + 3.75p = 1.25 + 2.5p$. The expected payoff from choosing Top (with Player 2 always responding with Left) is 2.50. It is easy to see that for any $p>1/2$, Bottom dominates Top in expected payoff terms and Bottom could be part of an equilibrium strategy profile.

Working off the Geanakoplos et al. (1989) model, Dufwenberg and Gneezy (2000), using an experimental lost wallet game which has a similar structure to ours, explicitly measure the beliefs that players possess in this class of games. They find, in this paired game with sequential moves, that the amount of money that the second player allocates between herself and the first mover depends crucially on the second mover's expectations about the first mover's expectations about this amount of money. Thus Dufwenberg and Gneezy (2000) provide strong corroboration of the Geanakoplos et al. proposition that players' beliefs play a pivotal role in such games. We, however, did not explicitly measure players' beliefs in our study and thus cannot directly draw any conclusions about the nature of such beliefs – a lapse that we intend to remedy in future studies.

Before moving on to discussing the main insights coming out of this study we would like to briefly discuss prior studies which have looked at the role of trust in social dilemmas using the SVO scale. Parks et al. (1995, 1996) also uses the Yamagishi SVO scale to look for correlation between people's trust score and their behavior in social dilemma situations. Parks and Hulbert (1995) investigate the effects of trust on subjects' response to fear in a social dilemma. They find (among other things) that high trusters cooperate more than low trusters when fear was present but cooperate at the same rate when fear is absent. Parks, Henager and Scamahorn (1996) look at high and low trusters' responses to messages of intent in a prisoner's dilemma game. They find that low trusters respond to competitive messages by reducing cooperation but were unaffected by cooperative messages. High trusters on the other hand react to cooperative messages by increasing cooperation but are unaffected by competitive messages. Parks (1994) finds that a subject's score on the Yamagishi trust scale is a good predictor of that subject's contributions to a public goods game. Using an expanded set of questions (which include the SVO questions), Yamagishi and Sato (1986) examine the role of fear and/or greed in an experimental public goods game where the public good can be provided in a conjunctive, disjunctive or additive manner. Half the subjects played with friends who knew each other and the other half interacted with strangers who were in separate compartments and did not have a chance to see or talk to one another. One of the hypotheses of the study is that greater mutual trust among friends would make them contribute more to the public goods game in the conjunctive condition. The authors find support for this hypothesis when the size of the public good was relatively large.

There are three main insights arising from this current study. First, people often deviate from game theoretic predictions by exhibiting notions of trust and reciprocity. As Rabin (1998, p. 16-17) says,

> "…probably no other two-word description of human motives comes close to "self-interest" in explaining economic behavior. Yet, pure self-interest is far from a complete description of human motivations and realism suggests economists should move away from the presumption that people are *solely* self-interested."

Second, and more importantly, trust and trustworthiness are not correlated. In fact trust and trustworthiness are fundamentally different constructs. See Chaudhuri, Sopher and Strand (2002) for further elaboration of this point. What has been interpreted as trusting behavior in prior studies (such as Berg, Dickhaut and McCabe, 1995 or Gneezy et al. 2000) has two distinct components. One is "pure trust" while the other is an element of calculated risk taking or a predilection for accepting a gamble. The former component is definitely a "social virtue" (as defined by Fukuyama, 1995), the latter probably not.

Finally and equally importantly, we find that when it comes to the idea of social capital (See Putnam, 2000 or Fukuyama, 1995)[14] it is trustworthiness that is more important and relevant rather than trust. Trustworthiness implies trust but the converse is not true. That is if one is trustworthy, then one is definitely trusting but a trusting individual is not necessarily trustworthy.

[14] Putnam (2000, Chapter 8, p. 136-7) comments "Other things being equal, people who trust their fellow citizens volunteer more often, contribute to charity, participate more often in politics and community organizations, serve more readily on juries, give blood more frequently, comply more fully with their tax obligations, are more tolerant of minority views, and display many other forms of civic virtue." A growing body of research suggests that "social capital" as embodied in the tendencies to "trust" and to "reciprocate" the trust of strangers influence a wide range of economic phenomena and activities. See for instance Fukuyama (1995), La Porta et al. (1997) and Knack and Keefer (1997).

Dawes and Thaler (1988, p. 195) provide an anecdote that nicely sums up this insight that some players are consistent in their trust and reciprocity while others are more opportunistic.

> "In the rural areas around Ithaca it is common for farmers to put some fresh produce on the table by the road. There is a cash-box on the table, and customers are expected to put money in the box in return for vegetables they take. The box has just a small slit, so money can only be put in, not taken out. Also, the box is attached to the table, so no one can (easily) make off with the money. We think that the farmers have just about the right model of human nature. They feel that enough people will volunteer to pay for the fresh corn to make it worthwhile to put it out there. The farmers also know that if it were easy enough to take the money, someone would do so."

4. Conclusion

Our first hypothesis, which predicted that we will see a large number of {Bottom, Right} outcomes, despite the fact that {Top, Left} is the SPE outcome, is vindicated. This is in accordance with previous studies and the fact that people are motivated by trust and reciprocity to deviate from the SPE outcome predicted by game theory. Our second hypothesis, which predicted a correlation between trust scores and trusting and/or reciprocal moves is partially rejected. However, we do find strong support for a modified and subtler version of this hypothesis. We find two different types of players - consistent and inconsistent - whose behavior was correlated with their trust scores. Consistent players, who both trusted as Player 1 and reciprocated as Player 2, were characterized by statistically higher trust scores than the inconsistent players. The consistent players are more likely to be driven by their pure trust levels. Inconsistent players had statistically lower trust scores and were possibly more interested in taking advantage of the trust and trustworthiness of others in maximizing their own payoff at the expense of achieving payoff equity at the SJPM outcome. We see these latter decisions as being independent of

people's trust levels and may have more to do with taking calculated risks of expecting trust and reciprocity from others, while being unwilling themselves to show reciprocity.

Our experiment situates itself within a body of work questioning the validity of economic predictions based on self-interest and profit-maximization. Further studies in the dynamics of trust and reciprocity could vastly expand our understanding of real-world economic behavior. A greater understanding of such emotional dispositions could have important repercussions in designing incentive contracts, negotiations, community development, and in framing policy.

References

Abbink, K., Irlenbusch, B., and Renner, E., (2000), "The Moonlighting Game: An Experimental Study on Reciprocity and Retribution", *Journal of Economic Behavior and Organization*, 42, p. 265-277.

Berg, J., J. Dickhaut and K. McCabe, (1995), "Trust, Reciprocity and Social History", *Games and Economic Behavior*, 10, p. 122-42.

Camerer, C. (2001), "Social Preferences in Dictator, Ultimatum and Trust Games", Chapter 3 in "Behavioral Game Theory: Experiments on Strategic Interaction", Mimeo, California Institute of Technology, forthcoming from Princeton University Press.

Chaudhuri, A., Sopher, B., and Strand, P. (2002), "Cooperation in Social Dilemmas, Trust and Reciprocity", *Journal of Economic Psychology*, 23(2), p. 231-249

Chaudhuri, A. and L. Gangadharan (2002), "Gender Differences in Trust and Reciprocity", Working Paper No. 2002-03, Department of Economics, Wellesley College, Wellesley, MA

Cox, J. (2002), "Trust, Reciprocity and Other-Regarding Preferences: Groups vs. Individuals and Males vs. Females", Chapter 14, in Rami Zwick and Amnon Rapoport (Eds.), Experimental Business Research, Kluwer Academic Publishers, Norwell, MA and Dordrecht, The Netherlands.

Dufwenberg, M. and U. Gneezy (2000), "Measuring Beliefs in an Experimental Lost Wallet Game", *Games and Economic Behavior*, 30, p. 163-182.

Dawes, R. and R. Thaler (1988), "Anomalies: Cooperation", *Journal of Economic Perspectives*, 2(3), Summer, p. 187-97.

Eckel, C. and R. Wilson (2002), "Conditional Trust: Sex, Race and Facial Expression in a Trust Game, Paper presented at the 2002 Annual American Economic Association Meetings, Atlanta, GA.

Fukuyama, F. (1995), "Trust: The Social Virtues and the Creation of Prosperity", The Free Press, New York, NY.

Geanakoplos, J., D. Pearce and E. Stacchetti, "Psychological Games and Sequential Rationality", *Games and Economic Behavior*, 1, p. 60-79.

Glaeser E., D. Laibson, J. Scheinkman and C. Soutter, (2000), "Measuring Trust", *Quarterly Journal of Economics*, 115(3) August 2000, p. 811-846.

Gneezy, U., Güth, W., and Verboven, F., (2000), "Presents or Investments? An experimental analysis", *Journal of Economic Psychology*, 21, p. 481-493.

Knack, S. and P. Keefer (1997), "Does Social Capital Have an Economic payoff? A Cross-Country Investigation", *Quarterly Journal of Economics*, 112(4), p. 1251-88.

Kramer, R. (1999), "Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions" Annual Review of Psychology, Palo Alto, 1999.

Kuhlman, D. M., C. R. Camac, and D. A. Cunha (1986), "Individual Differences in Social Orientation", in H. Wilke, D. M. Messick and C. G. Rutte (Eds.) Experimental Social Dilemmas, p. 151-176, Frankfurt, Peter Lang.

La Porta R., F. Lopez-de-Silanez, A. Shleifer and R. Vishny (1997), "Trust in Large Organizations", *American Economic Review Papers and Proceedings*, 87(2), p. 333-338.

McCabe, K. M. Rigdon and V. Smith (2002), "Cooperation in Single Play, Two-Person Extensive Form Games Between Anonymously Matched Players", Chapter 3 in Rami Zwick and Amnon Rapoport (Eds.), Experimental Business Research, Kluwer Academic Publishers, Norwell, MA and Dordrecht, The Netherlands.

McCabe, K., S. Rassenti and V. Smith (1998), "Reciprocity, Trust and Payoff Privacy in Extensive Form Bargaining", *Games and Economic Behavior*, 24, p. 10-24.

Messick, D. M. and C. G. McClintock (1968), "Motivational Bases of Choice in Experimental Games", *Journal of Personality and Social Psychology*, 4, p. 1 – 25.

Parks, C.D. (1994), “The Predictive Ability of Social Values in Resource Dilemmas and Public Goods Games”, *Personality and Social Psychology Bulletin*, 20(4), p. 431-38.

Parks, C.D. and L. G. Hulbert, (1995), “High and Low Trusters’ Responses to Fear in a Payoff Matrix”, *Journal of Conflict Resolution*, 39(4), p. 718-30.

Parks, C.D., R.F. Henager and S.D. Scamahorn, (1996), “Trust and Reactions to Messages of Intent in Social Dilemmas”, *Journal of Conflict Resolution*, 40(1), p. 134-51

Putnam, R.D. (2000), “Bowling Alone: The Collapse and Revival of American Community”, Touchstone, New York, NY, 2000.

Rabin, M. (1998), “Psychology and Economics”, *Journal of Economic Literature*, XXXVI, p. 11 – 46.

Schotter, A., Weigelt, K., and Wilson, C., (1994), “A Laboratory Investigation of Multiperson Rationality and Presentation Effects”, *Games and Economic Behavior*, 6, p. 445-468

Yamagishi, T., (1986), "The Provision of Sanctioning as a Public Good", *Journal of Personality and Social Psychology,* 51 (1), p. 110-116.

Yamagishi, T. and K. Sato, (1986), "Motivational Bases of the Public Goods Problem", *Journal of Personality and Social Psychology,* 50(1), p. 67-73.

Yukawa, M. (1985), “Structural and Psychological Factors in Social Dilemmas”, Unpublished Bachelor’s Thesis, Hokkaido University, Sapporo, Japan.

Appendix

SUBJECT ID#_____

Instructions

This is an experiment in the economics of market decision making. Various funding agencies including Wellesley College have provided funds to conduct this research. The instructions are simple. If you follow them closely and make appropriate decisions, you may make an appreciable amount of money. All payoffs are denoted in dollars and cents. These earnings will be paid to you in cash at the end of the experiment. The experiment consists of two parts.

We will carry out Part 1 first and then proceed to Part 2.

Part 1:

The first part of this experiment consists of filling out a short 5-question survey. Please answer the survey by circling the appropriate response to each question.

Age: _____years Subject ID # _____

Class Year: _____

For each of the five questions below, please ***circle*** one of the five options given.

Question 1. Most people tell a lie when they can benefit by doing so.

Strongly Disagree	Mildly Disagree	Neutral	Mildly Agree	Strongly Agree

Question 2. Those devoted to unselfish causes are often exploited by others.

Strongly Disagree	Mildly Disagree	Neutral	Mildly Agree	Strongly Agree

Questions 3. Some people do not cooperate because they pursue only their own short-term self-interest. Thus, things that can be done well if people cooperate often fail because of these people.

Strongly Disagree	Mildly Disagree	Neutral	Mildly Agree	Strongly Agree

Question 4. Most people are basically honest.

Strongly Disagree	Mildly Disagree	Neutral	Mildly Agree	Strongly Agree

Question 5. One should not trust others until one knows them well.

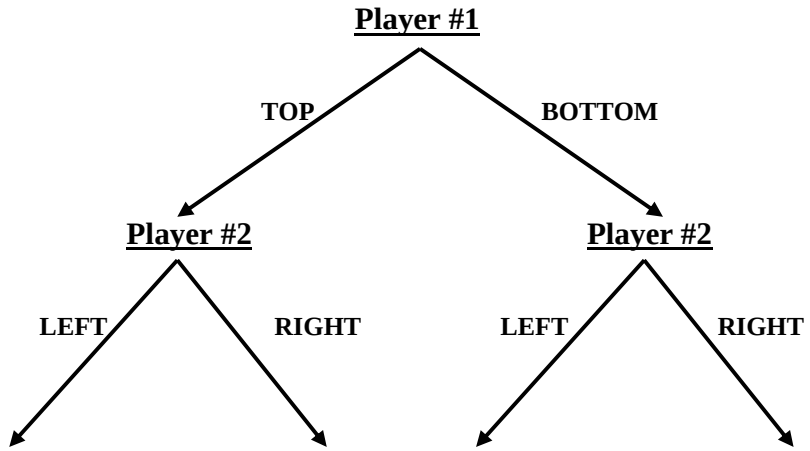
Strongly Disagree	Mildly Disagree	Neutral	Mildly Agree	Strongly Agree

---

Part 2:

The second part of the experiment consists of a decision making game. This experiment will be conducted in pairs. Each of you will always be paired with another person who will be in the other room and you will not be told of the identity of your partner at any point in the game.

This is a sequential game, with one player designated as Player 1 and the other as Player 2. To start the game, Player 1 will choose between two strategies: Top and Bottom. Player 2 will get to see what Player 1 chose and will respond by picking between the strategies: Left and Right. The payoff to each player depends on the choices made by both. The payoffs are denoted in dollars and cents and will be paid to you in cash at the end of the experiment. To understand how the game works please look at the diagram below.



Payoffs:				
PLAYER 1 →	**2.50 $**	**0.75 $**	**1.25 $**	**3.75 $**
PLAYER 2 →	**1.25 $**	**0.75 $**	**5 $**	**3.75 $**

SUBJECT ID#_____

If Player 1 chooses Top and Player 2 chooses Left then Player 1 gets $2.50 and Player 2 gets $1.25. If Player 1 chooses Top and Player 2 chooses Right, then Player 1 gets $0.75 and Player 2 gets $0.75. If Player 1 chooses Bottom and Player 2 chooses Left then Player 1 gets $1.25 and Player 2 gets $5.00. If Player 1 chooses Bottom and Player 2 chooses Right then Player 1 gets $3.75 and Player 2 gets $3.75.

Each of you will play both roles in this experiment. You will not be paired with the same person twice. In one pair you will be Player 1 while in the other pair you will be the Player 2.

Decisions made by each player will be conveyed using the Decision Sheet. Please take a look at the decision sheet now. You as, Player 1 will first record your decision on the decision sheet. This will then be collected by the experimenter and taken to the other room and given to the Player 2 you are paired with. The Decision Sheet will be brought back to you after Player 2 has made a decision.

Each player will make a Player 1 decision first. This will be followed by each person making a Player 2 decision.

You will record your payoffs in the Record Sheet. Please take a look at the Record Sheet now. When you are Player 1, you will enter your Player 1 decision in the appropriate box. Then, after you are informed of what the Player 2 you are paired with has decided, please record your payoff in the appropriate box. When you are Player 2, you will be informed of what the Player 1 you are paired with has decided. You will then make your decision and enter the corresponding payoff in the appropriate box on the Record Sheet.

Record Sheet **SUBJECT ID#_____**

AS PLAYER 1		
I CHOSE	PLAYER 2 CHOSE	PAYOFF
AS PLAYER 2		
PLAYER 1 CHOSE	I CHOSE	PAYOFF
	TOTAL PAYOFF =	

Decision Sheet

Player #1 (Subject ID # _____)
☐ **TOP**
☐ **BOTTOM**
(Check a box)
Player #2 (Subject ID #_____)
☐ **Left**
☐ **Right**