University of Nevada, Reno

Essays in Behavioral and Experimental Economics

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Economics

by.

Rattaphon Wuthisatian

Professor Mark Pingle, Ph.D. / Dissertation Advisor

May, 2015

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

RATTAPHON WUTHISATIAN

Entitled

Essays in Behavioral and Experimental Economics

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Mark Pingle, Ph.D., Advisor

Federico Guerrero, Ph.D., Committee Member

Mark Nichols, Ph.D., Committee Member

James Sundali, Ph.D., Committee Member

Thomas Quint, Ph.D., Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

May, 2015

Abstract

My dissertation, Essays in Behavioral and Experiment economics, is composed by three different research studies. In Chapter 1, the study offers an alternative theory of decision making under uncertainty. The presence of some perceived ambiguity together with some ambiguity tolerance leads a risk-averse decision-maker with diminishing marginal utility over all income levels to nonetheless take a chance. This alternative theory also provides an explanation for the equity premium puzzle, and for why a risk-averse buyer of insurance may also exhibit loss aversion. In Chapter 2, the study examines the effects of punishment, communication, and the interaction of these mechanisms on trusting and trustworthy behaviors in a laboratory experiment. The results suggest that the presence of both punishment and communication options places a downward pressure on the trustee to reduce trustworthy behaviors. However, if trustor expresses the highest level of trust accompanied by communication without implementing any punishment threat, the trustee's trustworthiness significantly increases. In Chapter 3, the study presents a theoretical model of play in the trust game that specifically features the potential which proposing equality has for promoting positive reciprocal behavior. The model proposes the utility function where the second mover appreciates the first mover that intends to split all earnings equally, and rewards this intention with positive reciprocity.

Table of Contents

List of Tables

List of Figures

Chapter 1

Taking Chances Versus Not: Ambiguity May Explain a Lot

Abstract

It is theoretically challenging to explain why the same person would rationally buy insurance but also gamble. Likewise, it is challenging to explain why the same person would take a chance to avoid a loss (i.e., exhibit loss aversion), yet also not take chances on comparable uncertain gains. The standard explanation has been a preference explanation: The decision-maker is risk seeking with increasing marginal utility over some income levels but risk averse with decreasing marginal utility over other income levels. Here, we provide an alternative explanation: The presence of some perceived ambiguity together with some ambiguity tolerance leads a risk-averse decision-maker with diminishing marginal utility over all income levels to nonetheless take a chance. This alternative theory also provides an explanation for the equity premium puzzle, and for why a risk-averse buyer of insurance may also exhibit loss aversion.

1. Introduction

When people make choices under uncertainty, they sometimes exhibit puzzling behaviors. A person who buys insurance also gambles. A person who takes a chance to avoid a loss also does not take a chance to obtain a comparable gain. One solution to these two puzzles is the hypothesis that the decision maker is risk averse in some situations but risk seeking in others, or equivalently that marginal utility is decreasing over some changes in wealth but increasing over others. Here, an alternative hypothesis is presented: Puzzling and seemingly contradictory behaviors under uncertainty can be explained by the ambiguity the decision maker perceives combined with the decision maker's degree of ambiguity tolerance, with the characteristics of the decision also playing a role.

Friedman and Savage (1948) show the coexistence of gambling and insurance can be explained by a utility function with convex as well as concave segments. Markowitz (1952) suggests a utility function with two concave segments and two convex segments. The arguments of Freidman, Savage, and Markowitz are well-known, yet scholars have continued to explore the shape of the utility function (e.g., Gregory (1980), Katz (1983), Dobbs (1988)). In offering their innovation, Friedman and Savage note that scholars preceding them had been reluctant to accept the assumption of increasing marginal utility, and their important contribution did not extinguish the reluctance. Their work has been followed by explanations of the coexistence of gambling and insurance that do not drop the concavity assumption, including transaction costs (Flemming, 1969), restrictions

on borrowing (Hakansson, 1970), and the indivisibility of expenditures (Ng, 1975). The innovation in this paper is to place ambiguity at the heart of an explanation.

In developing their prospect theory, Kahneman and Tversky (1979) followed the Friedman-Savage-Markowitz approach. A theory of loss aversion is introduced by assuming the utility function convex for negative changes in wealth, while a theory of risk aversion is maintained for gains by assuming the utility function concave for positive changes in wealth. More precisely, for gains, their theory implies a decision maker will tend to prefer a sure gain to a chance which yields that same gain as an expected value. Conversely, there will be a tendency to take a chance to avoid a sure loss equal to the expected value of the chance.

The assumption that the utility function is convex in losses can explain loss aversion, but it presents a problem for explaining why the same person would also buy insurance, so prospect theory needs an additional modification to describe behavior under uncertainty more generally. The standard modification (e.g., Kahneman and Tversky (1984, 1990), Kosegi and Rabin (2006), Barseghyan et al. (2011)) proposes that decision makers to do not apply objective probabilities but rather apply subjective probabilities that are biased in a specific way. The bias imbedded into the standard "probability weighting function" is that the decision-maker's subjective probability is greater than the objective probability when the probability is low and less than the objective probability when the probability is high. In the typical insurance situation, the objective probability of loss is low and the probability of no loss is high. Therefore, if a decision maker sufficiently overweighs the low probability, then the risk seeking behavior present in the

convex utility function can be overcome, so the decision maker should rationally to pay a premium for insurance to avoid taking a chance.

Like prospect theory, our theory assumes the decision maker does not apply an objective probability distribution with precision. However, rather than directly assuming a bias, our theory assumes the known distribution is hidden to some degree by perceived ambiguity. Once ambiguity is introduced, outcomes cannot be objectively weighted by a single probability distribution. The Hurwicz (1951) Criterion suggests people cope with ambiguity by giving decision weight to outcomes, ranging from the extremely optimist maxmax criterion to the extremely pessimistic maxmin criterion. Our theory is a derivative of the Alpha Minimax Expected Utility ($\alpha - MMEU$) theory, developed by Gilboa and Schmeidler (1989), a general theory of decision making under uncertain which combines expected utility theory with the Hurwicz Criterion.

In our version of the $\alpha - MMEU$ theory, the decision maker's degree of optimism is parameterized (as in the Hurwicz Criterion), and so is the degree of ambiguity. An underlying "anchor probability distribution" is the unique probability distribution the decision maker will apply when there is no ambiguity. This distribution is never subjectively biased by the decision-maker, and our utility function is always concave so the decision maker is always risk averse. However, the decision-maker's degree of optimism introduces a bias once ambiguity is present. Pessimism, or a lack of optimism, implies ambiguity only reinforces the assumed risk aversion and makes the decision maker not want to take changes. Optimism, on the other hand, can make ambiguity offset risk aversion. If there is enough optimism, then more ambiguity and a larger maximum

possible outcome (or stake) each offset the risk aversion, making it possible for the risk aversion decision maker to nonetheless rationally take a chance.

Because ambiguity is necessary for chance taking in our theory, it is worth recognizing that ambiguity is ubiquitous. In particular, previous work indicates people will tend to perceive ambiguity even when it might seem there should be none. Mulligan (2012) points out that our knowledge of external reality can never be perfect. We are cognitively limited, often lacking the sophistication necessary to assess probabilities and apply them.   Moreover, our knowledge that external reality can be biased can introduce ambiguity. For example, it is well known that Joseph Jagger, a mechanical engineer, is attributed "breaking the bank at Monte Carlo" in 1873 (even though he did not put the casino out of business) by identifying nine numbers on a particular roulette wheel which occurred more frequently.

Gambling fallacies may, at least in certain situations or to some extent, amount to decision maker optimism combined with perceived ambiguity. Walker (1985; 1992) provides evidence that gamblers generally hold a set of false beliefs about the likelihood of winning and the nature of gambling, with a tendency toward over-estimating their own ability to win. Langer (1975) finds that raffle ticket buyers generally believe the ticket they choose themselves is more likely to win than the ticket the vendor chooses. Similarly, Dickerson, Fabre and Bayliss (1986) reported that about 44% of TAB punters (horse race betters in New Zealand and Australia) reveal that their bet selections are based upon their own skills. Levy and Levy (2001) find decision-makers often construct their own subjective beliefs about winning a chance that are larger than objective

probabilities. These biases all have a flavor of optimism, and if ambiguity is not evidently present (as it is with horseracing) then part of a gambler's fallacy might be the perception of ambiguity when little to none is present.

Our theory also offers a solution to the equity premium puzzle. Mehra and Prescott (1985) demonstrated that the difference between the average rate of return on stock from holding bonds is too large to be explained by risk aversion alone. Several possible behavioral explanations to the puzzle have been offered, including fear of disaster (Reiz, 1988), habit information (Otrok, Ravikumar and Whiteman, 2002), first-order risk aversion (Epstein and Zin, 1990), and disappointment aversion (Ang, Bekart and Liu, 2000). The most notable explanation is "myopic loss aversion" (Benartzi and Thaler, 1995), the idea that stocks are especially avoided because they can generate short term losses. Our model suggests ambiguity will augment risk aversion in the typical stock purchase situation, more so than with bonds, offering an explanation for the premium return required for the purchase of stock.

The remainder of the paper is organized as follows. Section 2 provides a model of decision making under uncertainty, first examining no ambiguity, then total ambiguity, and then the $\alpha - MMEU$ model which combines risk and ambiguity. Section 3 derives the general implications of the model. Section 4 applies the model to four behaviors that have been of general interest: Section 5 concludes.

2. Model

Assume a decision-maker DM must choose an alternative a from N mutually exclusive alternatives $A = \{1,2,...,N\}$. DM perceives the outcome of action a depends

upon which state θ arises from among mutually K exclusive states $\Theta = \{1,2, \dots, K\}$.

When alternative a is chosen and state θ occurs, DM experiences the outcome $x = x(a, \theta)$, which provides utility $u(x(a, \theta))$. It is assumed $u'(x) > 0$, meaning DM perceives more is better. It is also assumed $u''(x) < 0$, meaning DM is risk averse.

2.1 Special Case 1: No Ambiguity

When there is no ambiguity, the probability of each possible state is known, so DM experiences pure risk. To be specific, assume the likelihood of state θ_i is $f(\theta_i) \in [0,1]$ and $\sum_{i=1}^{i=K} f(\theta_i) = 1$. In this case, expected utility theory asserts that an optimal choice by DM will maximize expected utility, where the expected utility of action a is given by

(1) $\quad E[u(x(a, \theta))] = \sum_{i=1}^{i=K} f(\theta_i) u(x(a, \theta)).$

To consider the simplest case, assume there are just two states of nature: good and bad. The good state occurs with probability p, while the bad state occurs with probability $1 - p$. There are two alternatives available to DM. One is the uncertain "prospect" (p, x_{min}, x_{max}), which yields the outcome x_{max} in the good state and $x_{min} < x_{max}$ in the bad state. The other alternative is a choice that yields the expected outcome

(2) $E[x] \equiv \bar{x} = px_{max} + (1 - p)x_{min}$

with certainty.

The utility of the certain choice is $u(\bar{x}) = u(px_{max} + (1 - p)x_{min})$, the utility of the expected outcome. The utility of the prospect (p, x_{min}, x_{max}) is the expected utility

(3) $E[u(x)] = p\,u(x_{max}) + (1 - p)u(x_{min}).$

Figure 1 compares these to alternative choices. The well-known result is shown: If a decision-maker is risk averse, then the utility of receiving the expected outcome $\bar{x}$ with certainty is greater than the expected utility of the prospect (p, x_{min}, x_{max}); i.e., $u(\bar{x}) > E[u(x)]$.

Consequently, as shown in Figure 1, there exists a certainty equivalent $\hat{x}$, such that $u(\hat{x}) = E[u(x)]$. If DM repeatedly chooses the prospect, or if many decision makers like DM were to choose the prospect, then the mean outcome will be $\bar{x}$, which is greater than $\hat{x}$. This provides an opportunity for an organization to pool the outcomes of decision makers choosing the uncertain action and provide insurance. The quantity $\bar{x} - \hat{x}$ is the maximum amount the pooling organization (e.g. investment fund, insurance company) can take from each decision maker as a fee such that decision makers identical to DM would desire to trade away the uncertain prospect and receive in exchange a certain outcome provided by the pooling organization.

Figure 1: A Risk Averse Decision-Maker is Willing to Pay to Avoid a Risk
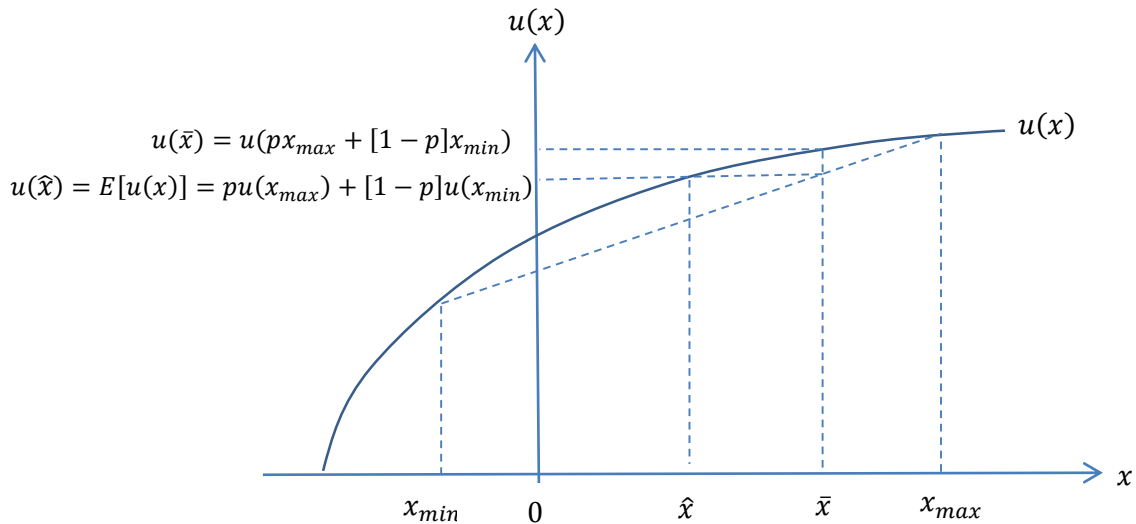
This standard case of decision making under pure risk provides a baseline against which decision making under ambiguity can be compared. A risk averse decision maker in a pure risk environment will not prefer the uncertain prospect (p, x_{min}, x_{max}) to receiving the certain outcome $\bar{x}$. Our interest is in identifying the circumstances in which a risk averse decision maker may nonetheless prefer the uncertain alternative when the uncertainty is not pure risk but involves some degree of ambiguity.

2.2 Special Case 2: Total Ambiguity

When DM faces total ambiguity, the choice must be made without any information about the probabilities of the possible states of nature. Facing this situation, the Hurwicz (1951) Criterion asserts DM will evaluate the uncertain prospect based upon the sizes of the outcomes. The simplest Hurwicz Criterion only allocates decision weight to the largest and smallest outcomes expected across the various possible states. DM is more optimistic when more weight is allocated to the largest outcome, and more pessimistic when more weight is allocated to the smallest outcome. Optimism is equivalent to being tolerant of the ambiguity, while pessimism is equivalent to being ambiguity intolerant.

Letting α denote the degree of optimism, or degree of ambiguity tolerance, the Hurwicz valuation of the uncertain outcome can be presented as

$$(4) \quad H = \alpha \left[u \left(\max_{\theta}[x(a,\theta)] \right) \right] + [1 - \alpha] \left[u \left(\min_{\theta}[x(a,\theta)] \right) \right], \qquad 0 \leq \alpha \leq 1.$$

When $\alpha = 1$, the Hurwicz Criterion becomes the maxmax criterion, and DM possesses the highest possible degree of optimism or ambiguity tolerance . When $\alpha = 0$, the

Hurwicz Criterion becomes the maxmin criterion, and DM possesses the highest possible degree of pessimism or ambiguity intolerance.

In the simplified two state of the world, the choice under uncertainty yields one of two outcomes. Call the two possible outcomes x_{min} and x_{max}, and say DM takes the "chance" (x_{max}, x_{min}) when the uncertain alternative is chosen. The Hurwitz criterion (4) indicates DM would in this case receive the utility

(5) $H = \alpha u(x_{max}) + [1 - \alpha]u(x_{min}).$

Why would risk averse DM choose to take the chance (x_{max}, x_{min}) when facing ambiguity rather than choose the alternative that yields $\bar{x}$ with certainty? The reason is optimism. Optimism plays no role in decision making when there is no ambiguity, but it can move DM to take the chance when the decision environment is ambiguous. More optimism places more decision weight on the maximum possible outcome in condition (5). This is what makes it possible for the chance to look better than the certain outcome.

2.3 Combining Risk and Ambiguity: An $\alpha - MMEU$ Model

The Alpha Minimax Expected Utility ($\alpha - MMEU$) model, developed by Gilboa and Schmeidler (1989), combines risk and ambiguity. The model used here, developed by Melkonyan and Pingle (2010), is a version of the $\alpha - MMEU$ which parameterizes the degree of ambiguity. This parameterization is useful because tradeoffs can be characterized that exist between the degree of ambiguity and other factors. The utility value of uncertain alternative a, given by our $\alpha - MMEU$ model, is given by:

(6) $V(a) = (1 - \lambda) \sum_{i=1}^{i=K} f(\theta_i) u(x(a, \theta)) + \lambda \left[\alpha \left[u \left(\max_{\theta} [x(a, \theta)] \right) \right] + [1 - \alpha] \left[u \left(\min_{\theta} [x(a, \theta)] \right) \right] \right]$

The parameter λ measures the degree of ambiguity, where $0 \leq \lambda \leq 1$. When $\lambda = 0$, the $\alpha - MMEU$ model reduces to the Expected Utility Model (1) and DM faces pure risk. Alternatively, when $\lambda = 1$, the model reduces to the Hurwicz Model (5) and DM faces total ambiguity. The general case lies between the extremes. The probability distribution $f(\theta)$ is called the "anchor probability distribution." When there is some ambiguity ($\lambda > 0$), the uncertainty is anchored by the distribution $f(\theta)$, but DM cannot see $f(\theta)$ with clarity. As the degree of ambiguity decreases, the distribution $f(\theta)$ becomes more apparent.

In the simple two state case, the $\alpha - MMEU$ model reduces to

(7) $\quad V(x) = [1 - \lambda][pu(x_{max}) + [1 - p]u(x_{min})] + \lambda[\alpha u(x_{max}) + [1 - \alpha]u(x_{min})].$

3. Implications

In this section, a series of meaningful implications of the $\alpha - MMEU$ model (7) are derived.

Proposition 1: For the chance (x_{max}, x_{min}), which yields minimax expected utility $V(x)$ given in (7) when there is ambiguity (i. e., $\lambda > 0$) and expected utility $E[u(x)]$ given in (3) when there is no ambiguity (i. e., $\lambda = 0$),

(8) $\quad V(x) \left(\genfrac{}{}{0pt}{}{>}{\genfrac{}{}{0pt}{}{=}{<}}\right) E[u(x)] \iff \lambda[\alpha - p][u(x_{max}) - u(x_{min})] \left(\genfrac{}{}{0pt}{}{>}{\genfrac{}{}{0pt}{}{=}{<}}\right) 0.$

Proof: See Appendix.

Proposition 1 tells us when the minimax expected utility $V(x)$ is greater than the expected utility valuation $E[u(x)]$ versus when it is less. That is, it tells us when ambiguity makes the uncertain alternative look better and when worse.

Proposition 2: If there is some ambiguity, so $\lambda > 0$, then $V(x) > E[u(x)]$ if and only if $\alpha > p$.

Proof: This proposition follows directly from Proposition 1 since $u(x_{max}) > u(x_{min})$.

Proposition 2 indicates DM must be sufficiently optimistic in order for ambiguity to enhance DM perception of the uncertain alternative. The sufficient level is greater when the likelihood of the good outcome is higher. Put another way, if DM is ambiguity intolerant (low α), then DM ambiguity will make the uncertain alternative look less attractive unless the likelihood of the good outcome is very low.

Condition (8) indicates the degree of ambiguity λ and the size of the stake $u(x_{max}) - u(x_{min})$ each magnify the perception DM has of the uncertain alternative. When DM is sufficiently optimistic, so $\alpha > p$, an increase in ambiguity λ or in the stake $u(x_{max}) - u(x_{min})$ increases the difference between $V(x)$ and $E[u(x)]$, providing more reason for DM to take the chance. Alternatively, when DM is sufficiently pessimistic, so $\alpha > p$, more ambiguity or larger stakes only reinforce not taking the chance.

Propositions 1 and 2 provide insight for empirical studies of ambiguity. Presuming people differ with regard to their degree of ambiguity tolerance (different levels for α), Proposition 2 indicates we would not expect subjects to reveal their different tolerances when the probability p of the good outcome is high. Rather, differences might be revealed when p is relatively low, so $\alpha > p$ for most subjects.

Because $\alpha > p$ is necessary but not sufficient for $V(x) > U(\bar{x})$, increasing the degree of ambiguity and the stake in the experiment should elicit responses that allow differing levels of ambiguity tolerance to be identified. Optimistic decision makers have high ambiguity tolerance (high α), so they will opt for chance alternatives over certain alternatives before less optimistic decision makers who have low ambiguity tolerance (low α).

Proposition 3: For the chance alternative yielding the minimax expected utility $V(x)$ given by (7) when there is ambiguity and expected outcome $\bar{x} = px_{max} + [1-p]x_{min}$ when there is no ambiguity,

(9) $\quad V(x) \begin{pmatrix} > \\ = \\ < \end{pmatrix} u(\bar{x}) \iff \lambda[\alpha - p] \begin{pmatrix} > \\ = \\ < \end{pmatrix} \dfrac{u(px_{max}+[1-p]x_{min})-[pu(x_{max})+[1-p]u(x_{min})]}{[u(x_{max})-u(x_{min})]} > 0.$

Proof: See Appendix

Proposition (3) indicates when risk aversion is overcome by optimism under ambiguity, and when it is not. Equivalently, condition (9) indicates when DM will opt to take the chance (x_{min}, x_{max}) over accepting the outcome $\bar{x}$ with certainty, and when not. Notice $\alpha > p$ is necessary for $V(x) > u(\bar{x})$, so will opt DM to take a chance, but not sufficient. It is apparent from condition (9) that DM will take with low degrees of ambiguity and ambiguity tolerance only if the probability p is very low. Proposition 4 more carefully delineates how the willingness of DM to take a chance depends upon the likelihood of the good outcome, the degree of ambiguity, and the degree of ambiguity tolerance.

Proposition 4: If there is some ambiguity (i.e., $\lambda > 0$), and some ambiguity tolerance

(i.e., $\alpha > 0$), then there exists a unique value $\bar{p}$ such that

(10) $\quad p \begin{pmatrix} < \\ = \\ > \end{pmatrix} \bar{p} \Leftrightarrow V(x) \begin{pmatrix} > \\ = \\ < \end{pmatrix} u(\bar{x}),$

and $\bar{p}$ is increasing in λ and in α, and $\bar{p} \to 1$ as $\lambda\alpha \to 1$

Proof: See Appendix

Figure 2 provides some intuition of Proposition 4, and offers insight into how the

three parameters p, λ, and α affect the willingness of DM to take the chance.   The two

sides of the right most inequality in (9) are plotted as they depend upon the probability p,

along with the function $f(p) = V(x) - u(\bar{x})$. The left side of (9) is linear and decreasing

in p, while the risk aversion assumption implies the right side increases from zero in p

but then deceases back to zero.   As p $\to$ 0, the left side of (9) converges to $\lambda\alpha$, while the

right side of (9) converges to zero. Thus, for values of p near zero, $V(x) > u(\bar{x})$.

Alternatively, as p $\to$ 1, the left side of (9) converges to $\lambda[\alpha - 1]$, while the right side of

(9) converges to zero. Thus, for values of p near 1, $V(x) < u(\bar{x})$.   Thus, we know DM

will take a chance under ambiguity when p is very low, and we know DM will not take a

chance when p is very high.

Figure 2: When Will Risk Averse DM Take a Chance Under Ambiguity?
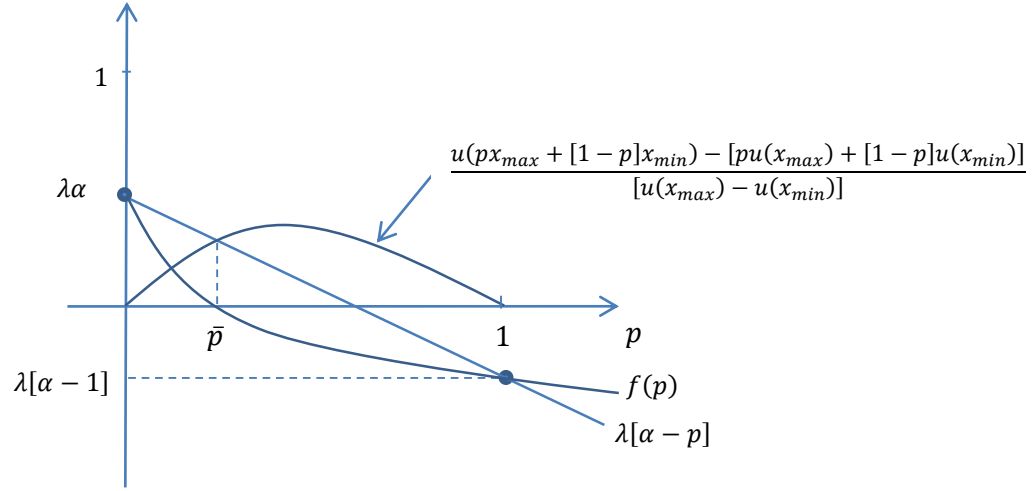
At $p = 0$, the derivative $f'(p)$ is negative, indicating an increase in p from zero makes taking the chance look less attractive. The case drawn in Figure 2 indicates $f(p)$ is strictly decreasing in p, which is one of two possible cases and is the case that will hold when the level of ambiguity is high. In the other case, f(p) decreases to a value less than $\lambda[\alpha - 1]$ and then increases to $\lambda[\alpha - 1]$ when $p = 1$. Because $f''(p)$ is positive over the whole domain, in either case, f(p) decreases until it equals zero at that value $p = \bar{p}$ and cannot return to zero, so $\bar{p}$ is unique. An increase in λ or α increase $\bar{p}$ up to a maximum value (less than one) when $\lambda\alpha = 1$. That is, when there is more ambiguity or more ambiguity tolerance (more optimism), it is more possible that DM will take chances when p is at a higher level.

Looking toward applications, condition (9) indicates, as long as $\alpha > p$, the amount of ambiguity λ can be small and the risk averse decision maker will still opt to take the chance when the stake $u(x_{max}) - u(x_{min})$ is large enough. A lottery is interesting in this respect because the stake gets larger as more people play. Also, the

probability of the good outcome, winning the lottery, gets smaller as more people play. Thus, even if there is only a small amount of ambiguity in the mind of the potential lottery player and just a small amount of optimism, then there is a lottery size large enough that all will want to play. We will show this carefully in the next section.

Similar logic can explain why people would rather tend to play a higher stakes bet, like playing a number in roulette rather than a color. We examine this below as a second example.

This model also offers an explanation for why the same person might be loss averse in one situation but buy insurance in another. We can examine a potential loss by making x_{min} a negative number (the possible loss) while making x_{max} equal to zero (no loss). For a given level of ambiguity tolerance α, our model indicates insurance will be desirable in the most typical situation where the probability of the loss is low or the probability p of no loss is high. Conversely, our model predicts loss aversion (i.e., take a chance in order to avoid a certain loss) in the opposite kinds of situations where p is low and the probability of loss is high. Our model also predicts optimistic people are more prone to displaying loss aversion. Insurance versus loss aversion is fleshed out below as a third application.

Finally, our model helps explain the equity premium puzzle, and this is presented as a fourth application. When the probability of the good outcome is reasonably high, which should be true for companies that attract investment in their equity securities, we would expect $\alpha < p$. With $\alpha < p$, more ambiguity and a higher stake compound the risk aversion of the decision maker in terms of making the equity investment look less

attractive relative to a sure thing. That is, the risk premium paid on equities is higher

than can be explained by risk aversion alone because ambiguity and high stakes increase

the premium required to make the investment alternative comparably attract to

alternatives with less ambiguity and lower stakes.

4. Application

Application 1: Why Might a Risk Averse Decision-Maker Play a Fair Lottery?

Consider a fair lottery in which each participant buys a dollar lottery ticket and

each ticket has an equal chance of winning. With N participants, the probability of

winning is $1/N$, and the probability of losing is $[N-1]/N$. The net winnings of the

winner is $N-1$ dollars, and the $N-1$ losers each lose one dollar. The expected value of

the lottery is therefore

(11) $\bar{x} = \frac{1}{N}[N-1] + \frac{N-1}{N}[-1],$

which is zero.   DM's expected utility associated with the purchase of a lottery ticket is

(12) $E[u(x)] = \left[\frac{1}{N}\right]u(N-1) + \left[\frac{N-1}{N}\right]u(-1),$

where x is the change in wealth experienced by DM, and $u(x)$ is the utility of this change

in wealth.

When there is no ambiguity ($\lambda = 0$), the assumed concavity of the utility function

implies the utility of the expected value of zero change in wealth that can be obtained by

not playing the lottery is strictly greater the expected utility of playing; $u\left(\frac{1}{N}[N-1] + \right.$

$\frac{N-1}{N}[-1]) = u(0) > \left[\frac{1}{N}\right]u(N-1) + \left[\frac{N-1}{N}\right]u(-1)$. That is, risk averse DM will not play the lottery if the uncertainty is pure risk, no matter how large the prize.

Friedman and Savage (1948) offer an explanation for why, under pure risk, someone would not only rationally play the lottery but also rationally buy insurance. They reference the following Alfred Marshall quote to provide the essence of their explanation: "Uncertainty … which does not appeal to great ambition or lofty aspirations has special attractions for very few … and acts as a deterrent to many. … But, if [it] offers a few extreme prizes, its attractiveness is increased all out of proportion…(Friedman and Savage 1948, p. 284) ." Friedman and Savage note that, prior to their analysis, economists were reluctant to assume anything other than a diminishing marginal utility for wealth. The contribution of Friedman and Savage (1948) was to show that one can explain gambling by assuming an increasing marginal utility for wealth at higher wealth levels, while maintaining a willingness to buy insurance by assuming diminishing marginal utility for losses. The behavioral intuition for assuming an increasing marginal utility at higher wealth level is the idea that people will seek risk when the increase in wealth is fundamentally significant, perhaps life changing. For example, Friedman and Savage mention moving up in socioeconomic class

Figure 3 illustrates the Friedman-Savage explanation for the lottery. Two situations are shown. There are more participants in situation B, so lottery prize $(N-1)_B$ is greater. In situation A, even though the marginal utility has started to increase, $E[u(x)]_A < u(0)$ holds, implying DM prefers not playing the lottery. The quantity $\hat{x}_A < 0$ is the maximum DM would be willing to pay to avoid playing the

lottery. In situation B, $E[u(x)]_B > u(0)$. The lottery prize $(N-1)_B$ is large enough

that the increasing marginal utility has sufficient impact to make playing the lottery

attractive. The quantity $\hat{x}_B > 0$ is the maximum DM would be willing to pay to play

the lottery, and it is this willingness to pay that allows organizations to offer lotteries that

are not fair but can still attract participants.

Figure 3: Friedman-Savage Utility Function
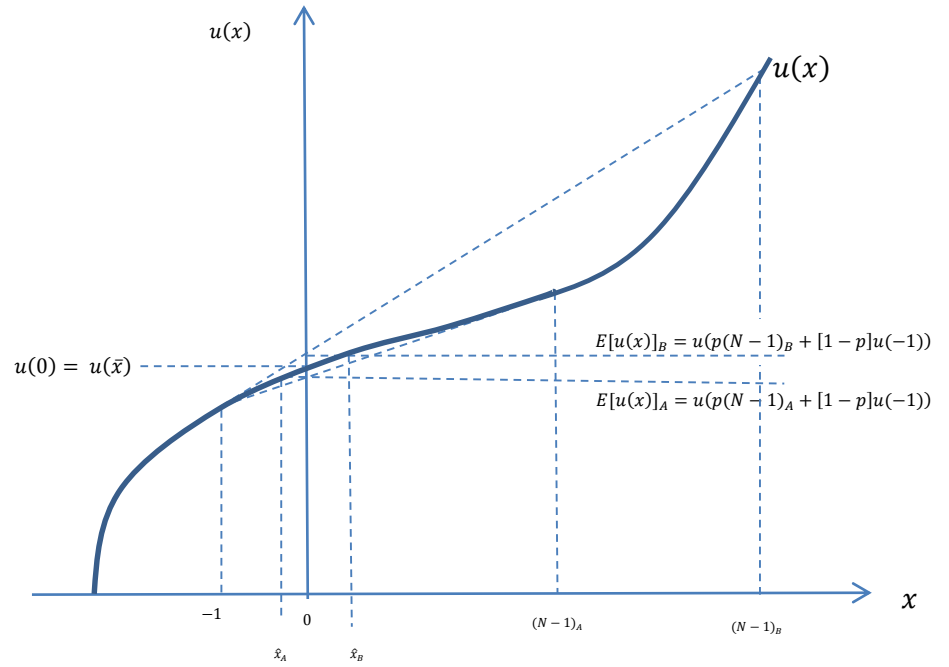
Our theory offers an alternative explanation to the Friedman-Savage explanation.

Rather than assuming the marginal utility of wealth increases at some point, our theory is

consistent with the thinking prior to Friedman and Savage which assumed the marginal

utility of wealth consistently diminishes. The willingness of DM to take a chance in our

theory does not rely upon risk seeking preferences, but rather relies upon a combination

of ambiguity and optimism.

Applying condition (7), the minimax utility of DM is

(13) $V(x) = [1 - \lambda]\left[\left[\frac{1}{N}\right]u(N-1) + \left[\frac{N-1}{N}\right]u(-1)\right] + \lambda\big[\alpha[u(N-1)] + [1-\alpha][u(-1)]\big].$

Because the expected value of this fair lottery is $\bar{x} = \frac{1}{N}[N-1] + \frac{N-1}{N}[-1] = 0$, the utility

of the expected value, $u(\bar{x}) = u(0)$, is the same as the utility of not playing the lottery. It

follows that DM will prefer playing the lottery when $V(x) > u(0)$. Using (9), $V(x) >$

$u(0)$ holds if and only if

(14) $\lambda\left[\alpha - \frac{1}{N}\right] > \dfrac{u(0) - \left[\frac{1}{N}u(N-1) + \frac{N-1}{N}u(-1)\right]}{u(N-1) - u(-1)}$

As the number of lottery participants N increases, the right side of (14) converges to zero

and the left side of (14) converges to $\lambda\alpha$. Therefore, as long as there is some ambiguity

$(\lambda > 0)$ and some ambiguity tolerance $(\alpha > 0)$, an increase in the number of participants

N will enhance the attractiveness of the lottery and eventually induce DM to play the

lottery.

Condition (14) cannot hold if $\alpha < 1/N$ or if λ is very small. Thus, our theory

indicates those attracted to the early stages of a lottery will be those who are more

optimistic (α large) and those who perceive more ambiguity (λ large). However, as the

size of the lottery prize grows, those who are more pessimistic and who perceive less

ambiguity will find the lottery attractive. Condition (14) indicates those who perceive

little ambiguity (λ near 0) and who are quite pessimistic (α near 0) will not tend to play

the lottery, but even these people will be attracted to play if the lottery prize gets large

enough.

Figure 4 presents how DM would view the lottery if there is no ambiguity. As the number of lottery participants increases, the expected value of the lottery remains at zero even though the prize increases from $(N-1)_A$ to $(N-1)_B$, However, the diminishing marginal utility assumption (or risk aversion assumption) causes expected utility to decrease from $E[u(x)]_A$ to $E[u(x)]_B$. Risk aversion also implies $E[u(x)]_A < u(0)$, so DM does not find the lottery attractive when the lottery prize is smaller. The increase in the size of the lottery makes the lottery look worse relative to not playing. DM would be willing to pay up to $\hat{x}_A$ to avoid playing the lottery when the lottery prize is $(N-1)_A$ and pay $\hat{x}_B$ to avoid playing when the lottery prize is $(N-1)_B$.

How do optimism and ambiguity combine to overcome this risk aversion and get DM to take a chance and play the lottery? When $\alpha > p = 1/N$, the optimism of DM places more weight on the good outcome (the utility $u(N-1)$ of the lottery prize) than the probability p puts on it when there is no ambiguity. More ambiguity magnifies this impact and allows the weight that optimism places on the good outcome to overcome the risk aversion. As the number of participants in the lottery grows, the stake $u(N-1) - u(-1)$ grows and the probability $p = 1/N$ decreases. Both of these changes also magnify the impact which the degree of optimism α places on the good outcome relative to the weight the probability p places on the good outcome. An extremely large lottery, can therefore attract a relatively optimist and risk averse decision maker.

Figure 4: The Lottery Under Pure Risk
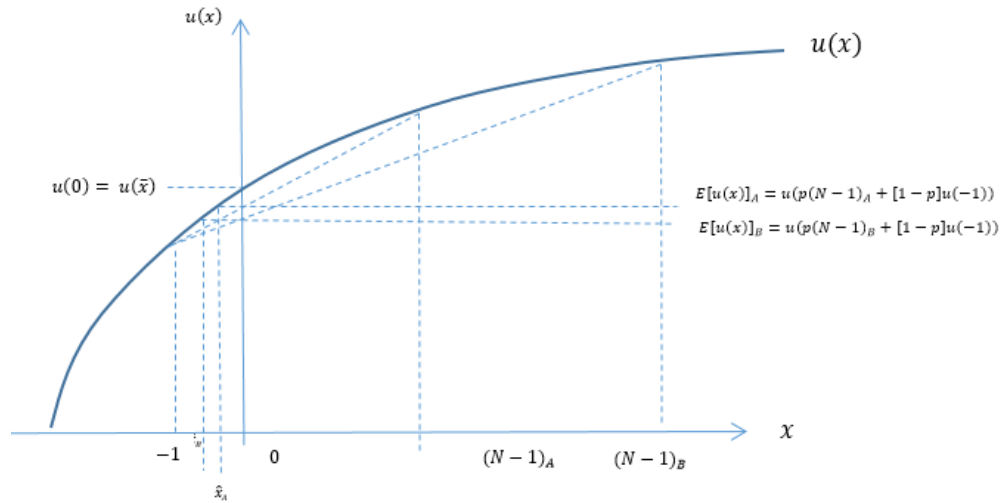
Application 2: Why might a Risk Averse Decision Maker Prefer a Higher Variance Bet?

The standard American roulette wheel has 38 slots; 18 Red, 18 black, and 2 green. The red and black slots are numbered 1 to 36. The green slots are numbered 0 and 00. A decision maker can make a variety of bets in the roulette game. Here, a $1 bet on red will be compared to $1 bet on a single number.

For a $1 bet on red, DM receives back the $1 bet plus an additional dollar if the ball falls into a red slot. Alternatively, DM loses the $1 bet amount if the ball falls into a black or green slot. Assuming the wheel is fair, the probability p of a red outcome is 18/38, implying the probability of not red is $1 - p = 20/38$. Thus, the expected outcome of the $1 bet on red is

$$(15) \quad \bar{x}_R = \frac{18}{38}[+1] + \frac{20}{38}[-1] = -\frac{2}{38} = -0.0526$$

For a bet on a single number, the payout is thirty-five to one. For example,

suppose DM bets \$1 on number 3. If the ball lands in the number 3 slot, DM receives the

\$1 bet back and an additional \$35. Alternatively, DM will lose the \$1 bet if the ball

lands in any of the other 37 slots. The probability p of winning \$35 is 1/38 while the

probability $1 - p$ of losing \$1 bet is 37/38, so the expected outcome of the \$1 bet on the

number 3 is

(16) $\bar{x}_3 = \frac{1}{38}[+35] + \frac{37}{38}[-1] = -\frac{2}{38} = -0.0526.$

Figure 5 presents how DM would perceive the two bets when there is no ambiguity.

For either bet, the expected outcome is the same: $\bar{x}_R = \bar{x}_3 = -0.0526$. Risk aversion

implies $E[u(\bar{x}_R)] < u(0)$ and $E[u(\bar{x}_3)] < u(0)$, so DM will neither find it attractive to

play the red bet nor the single number bet. The risk aversion assumption also implies

$E[u(\bar{x}_R)] > E[u(\bar{x}_3)]$, so the lower variance red bet appears better to DM than the higher

variance single number bet.

Figure 5: Under Ambiguity, A Risk Averse Decision-Maker is Willing to Play Roulette
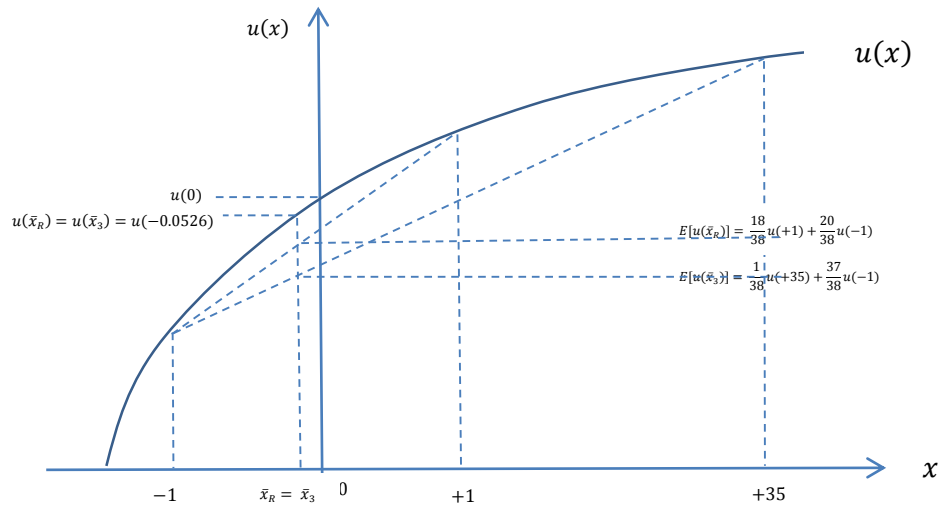
In reality, people not only play roulette, but they bet on single numbers much more than they bet on red. The Friedman-Savage assumption of increasing marginal utility can again explain this phenomenon, but our theory again offers an alternative.

Applying our condition (9), DM will prefer playing the bet on red to not playing at all when

$$(17) \quad \lambda \left[\alpha - \frac{18}{38} \right] > \frac{u(0) - [pu(+1) + [1-p]u(-1)]}{[u(+1) - u(-1)]}$$

and DM will prefer betting on a single number to not playing at all when

$$(18) \quad \lambda \left[\alpha - \frac{1}{38} \right] > \frac{u(0) - [pu(+35) + [1-p]u(-1)]}{[u(+35) - u(-1)]}.$$

Neither condition (17) nor (18) can hold if there is very little ambiguity (λ near 0), nor if the degree of ambiguity tolerance is very small (α near zero).   Thus, similar to the lottery case, our theory predicts that those who play roulette will be those who are more optimistic (α large enough) or perceive more ambiguity (λ large enough), but even more so a combination of the two.   For a given level of ambiguity, condition (18) will hold with a lower level of optimism α than condition (17). The valuation of the single number bet will therefore be higher. That is, if playing roulette is attractive for DM, the single number bet will be more attractive.

Application 3: Why a Risk Averse Decision Maker May Exhibit Loss Aversion Yet Still Buy Insurance

Suppose DM faces the chance of losing an amount of wealth equal to a, or a possible change in wealth of $-a$. Suppose the probability of experiencing no loss is p, so the probability of experiencing the loss is $1 - p$. Without insurance, the expected value of the loss is therefore

(19) $\bar{x} = p[0] + [1-p][-a] = -a[1-p] < 0$

The expected utility of DM is

(20) $E[u(x)] = pu(0) + [1-p]u(-a)$

    Figure 6 presents the standard explanation of how DM would value the loss in wealth relative to the purchase of insurance protection when there is no ambiguity. The diminishing marginal utility assumption (or the concavity of the utility function) implies $u(\bar{x}) > E[u(x)]$, so DM always prefers certainty and is willing to buy insurance in order to avoid the loss in wealth. Moreover, as shown in Figure 6, there exists a certainty equivalent $\hat{x}$ such that DM perceives the expected loss $-a[1-p]$ provides the same level of satisfaction as the certain payment $\hat{x}$. That is, by accepting the smaller certain change in wealth $-\hat{x}$, DM can avoid the risk of experiencing the larger change in wealth $-a$. The distance between $-a[1-p]$ and $\hat{x}$ is the "premium", or the maximum amount DM is willing to pay for insurance against the loss.

Figure 6: The Insurance Under Pure Risk
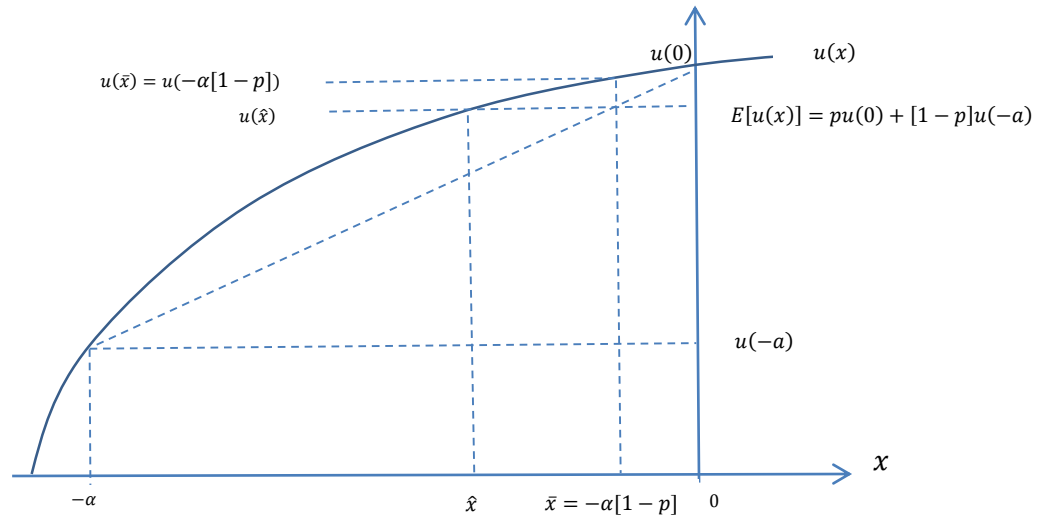
Our theory offers an explanation for why this risk averse decision maker may nonetheless choose to take the chance and self-insure. If DM takes the chance rather than accepting the sure loss of the expected outcome $\bar{x} = -\alpha[1 - p]$, then DM is exhibiting loss aversion. Thus, our theory offers an explanation of loss aversion that does not depend upon an increasing marginal utility of wealth.

Using condition (9), DM prefers taking the chance when $V(x) > u(\bar{x})$, or when

(21) $\lambda[\alpha - p] > \dfrac{u(-a[1-p])-[pu(0)+[1-p]u(-a)]}{[u(0)-u(a)]}$

The risk aversion assumption implies condition (21) will not hold if there is no ambiguity ($\lambda = 0$). When there is some ambiguity ($\lambda > 0$), the risk aversion assumption also implies condition (21) cannot hold if the degree of ambiguity tolerance α is small. In particular, condition (21) cannot hold if $\alpha < p$. In the typical insurance situation, the probability p of not incurring the loss is large, so $\alpha < p$ would be typical. Thus, our theory indicates DM would buy insurance in the typical situation where insurance is offered, for DM would have to be very optimistic (α near 1) to take the chance and not buy insurance.

Alternatively, DM will prefer to take the chance of losing the amount of wealth a when the probability p is small enough. We know of no other theory which can distinguish when the same person would buy insurance versus exhibit loss aversion and take the chance in a potential loss situation. Our theory does. With the degree of ambiguity determined by the situation and the degree of ambiguity tolerance determined by the decision maker's preferences, the probability $\bar{p}$ in condition (10) distinguishes the

situation when DM will want to buy insurance and mitigate risk from the situation when DM will be loss averse.   Proposition 4 indicates loss aversion will not occur when the probability of loss is high, unless there is much ambiguity and much ambiguity tolerance. DM will want to buy insurance in such situations.   More people will exhibit loss aversion when the probability $1 - p$ of the loss is high, or when the probability of no loss is low.   A higher stake will also tweak DM away from buying insurance and toward loss aversion

Given our theory, it is interestingly to revisit the now famous loss aversion example of Kahneman and Tversky (1981):

> Problem 1: Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
>
> If program A is adopted, 200 people will be saved.
>
> If program B is adopted, there is a 1/3 probability that 600 people will be saved and a 2/3 probability that nobody will be saved.
>
> Which of the two programs would you favor?

A substantial majority of subjects (72% of 152 subjects) chose program A. That is, subjects overwhelmingly exhibited risk aversion in this "gain frame," finding it more attractive to save 200 people with certainty than facing the risky prospect with an expected value equal to 200. For comparison, Kahneman and Tversky asked another group of subjects to choose between two programs in a "loss frame:"

Problem 2: Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

If program C is adopted, 400 people will die.

If program D is adopted, there is a one-1/3 probability that nobody will die and a 2/3 probability that 600 people will die.

Which of the two programs would you favor?

A substantial majority of subjects (78% of 155 subjects) chose program D. That is, subjects overwhelmingly exhibited risk seeking behavior in the loss frame. Analogous to the Friedman-Savage explanation for gambling behavior, the Kahneman-Tversky explanation for loss aversion is that the marginal utility of income is increasing in the domain where outcomes are negative. Our theory offers an alternative explanation.

Using our notation, $p = 1/3$ in the Kahneman-Tversky loss and gain frame examples. For the loss frame, this implies a 2/3 probability of loss, much higher than the typical insurance situation. In our model, this relatively low choice for p creates more potential for satisfying the $\alpha > p$ condition of Proposition 2, which is necessary for chance taking. For subjects with $\alpha > p$, our theory would indicate that the high stake of lost lives would further motivate chance taking, which in the loss frame would be loss aversion.

To illustrate the difference the probability and stake can make, suppose we increase the probability to the extreme of $p = 995/1000$, and suppose we change the

stake from 600 lives to a $600,000 home. The Kahneman-Tversky choice A would then become accept $597,000 with certainty. Choice B would become a lottery with a 995/1000 chance of receiving a $600,000 home and 5/1000 chance of receiving nothing. One would expect a substantial majority to select the certain choice A, as in the original Kahneman-Tversky experiment. Choice C would become losing $3,000 for sure (e.g., a "fair insurance premium), and choice D would become a lottery with a 995/1000 chance of losing nothing and a 5/1000 chance of losing the $600,000 home. The fact that so many chose to buy insurance in such situations is evidence that the loss aversion exhibited in the original Kahneman-Tversky experiment would NOT be exhibited.

Our model is challenged by the Kahneman-Tversky example in one interesting respect. With fixed values for λ and α, and single concave utility function like the isoelastic utility function, if DM is risk averse (risk seeking) in gains then DM will also be risk averse (risk seeking) in losses for the same value of p. That is, for our model to explain why DM might take a chance in a loss frame but not in a gain frame, with the utility function maintaining its concavity, λ or α must vary across the frames. While it is conceivable that DM could perceive more ambiguity in a loss frame than in the gain frame, it is more reasonable that the degree of optimism α might vary. Specifically, losses will loom larger than gains if DM is more optimistic (α larger) in the loss frame than in the gain frame.

To summarize, our theory suggests people will more likely exhibit loss aversion when the likelihood of loss is high but prefer buying insurance with the probability of loss is low. A higher level of perceived ambiguity and a higher degree of ambiguity

tolerance (optimism) accentuate this expectation. The same decision maker may rationally take chances in a loss frame but prefer certainty in a gain frame when potential losses make the decision maker more optimistic ("I will be one of the lucky ones") than potential gains.

Application 4: An Alternative Explanation of the Equity Premium Puzzle

Consider two investment opportunities, a stock share and a bond. If DM invests a particular amount in the stock, say $100, the probability of receiving the return x_{max} is p and the probability of receiving no return is $1 - p$. The expected return of the stock is therefore

(22) $\bar{x} = p\, x_{max} + [1 - p](0) = p\, x_{max},$

while the expected utility is

(23) $E[u(x)] = pu(x_{max}) + [1 - p]u(0).$

Assume DM can alternatively receive the outcome x_{bond} with certainty by investing the same amount in the bond.

The assumption of risk aversion implies the utility of the expected value of the stock is strictly greater than the expected utility; i.e., $u(\bar{x}) > E[u(x)]$. Thus, DM will not invest in the stock unless the expected outcome $\bar{x}$ is more than the certain outcome x_{bond}. If the entire market consists of investors identical to DM, then stock and bond investments can both occur only if $u(x_{bond}) = E[u(x)]$, so DM is indifferent between the two investments.   Thus, the relationship between x_{bond} and $\bar{x}$ must be as shown in Figure 7.   The difference $\bar{x} - x_{bond}$ is a measure of the equity premium received for bearing the risk associated with investing stock.

Figure 7: The Equity Premium Puzzle Under Pure Risk
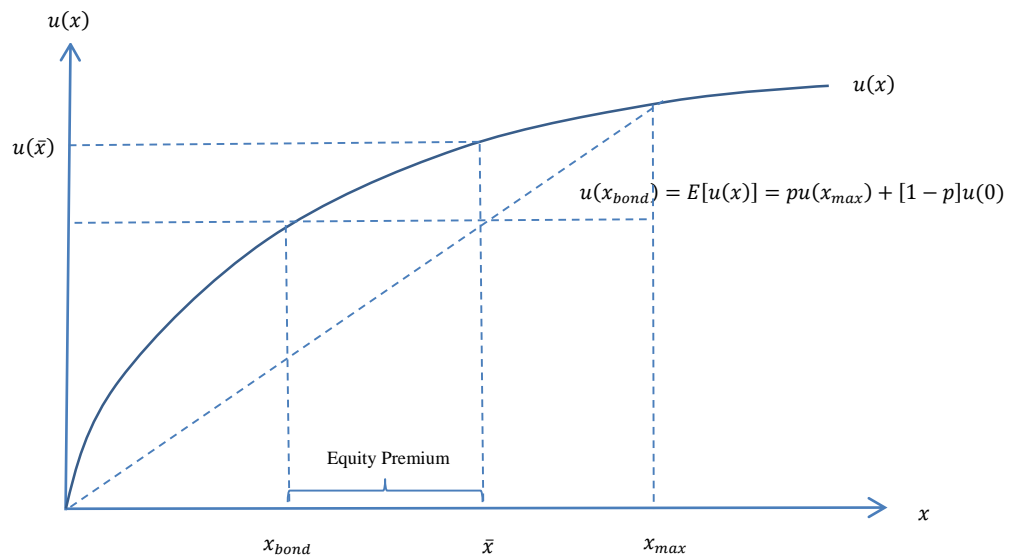
The equity premium actually observed in reality is larger than we would reasonably expect risk aversion to explain. According to Kocherlakota (1996), the estimated returns to investors from holding stocks were at 7% per year while the estimated returns from the U.S government bonds were at only one percent per year. Moreover, this phenomenon is regularly observed in every market around the globe. This unexpectedly large equity premium implies an inability of the risk aversion to reflect a realistic level of risk preferences among investors, and the challenge associated with explaining why the equity premium is so high is referred to as the "equity premium puzzle."

Several possible explanations to the equity premium puzzle have been proposed in economics and finance literature. "Loss aversion" seems to be the most widely recognized approach to rationalize the underlying behavior behind the puzzlingly large discrepancy between the returns of stocks and the fixed returns of bonds. A greater sensitivity to losses than to gains can explain a larger discrepancy than predicted by

standard expected utility theory, and a tendency to evaluate outcomes frequently also offers an explanation (Bernatzi and Thaler, 1995; Rabin and Thaler, 2001). Here, our theory offers an alternative explanation for why the equity premium might be larger than the amount $\bar{x} - x_{bond}$ shown in Figure 7.

When there is ambiguity ($\lambda > 0$), condition (8) implies $V(x) < E[u(x)]$ when $\lambda[\alpha - p][u(x_{max}) - u(0)] < 0$. Because companies issuing stock attract investment, it is reasonable to think that the anchor probability p is not extremely low, but rather might well be high. Consequently, for many decision-makers, even relatively optimistic decision makers, the probability p might well be high enough that $\alpha < p$. If $\alpha < p$, then $V(x) < E[u(x)]$. For DM to be indifferent between a stock investment with an uncertain expected return $\bar{x}$ and a bond investment with a certain return $x_{bond}^{ambiguity}$, $V(x) = u\left(x_{bond}^{ambiguity}\right)$ would have to hold, and $V(x) < E[u(x)]$ implies the equity premium $\bar{x} - x_{bond}^{ambiguity}$ must be greater than the equity premium $\bar{x} - x_{bond}$.

When $\alpha < p$, more ambiguity (λ larger) and a higher stake ($u(x_{max})$ is larger) will increase the difference $E[u(x)] - V(x)$. This would increase the equity premium needed to keep the uncertain stock investment equivalent to the certain bond investment. Thus, our theory indicates an exceptionally large premium would be expected when the uncertainty is especially ambiguous and when the stake is especially large.

To be more precise, since $x_{bond}^{ambiguity}$ is defined by $V(x) = u\left(x_{bond}^{ambiguity}\right)$ and

x_{bond} is defined by $E[u(x)] = u(x_{bond})$, it follows that $E[u(x)] - V(x) = u(x_{bond}) -$

$u\left(x_{bond}^{ambiguity}\right)$. Eliminating $E[u(x)] - V(x)$, we find

(24) $\lambda[p - \alpha][u(x_{max}) - u(x_{min})] = u(x_{bond}) - u\left(x_{bond}^{ambiguity}\right)$.

When $\alpha < p$, as hypothesized, each side of (24) is positive, which implies

$x_{bond}^{ambiguity} < x_{bond}$.

When we hold x_{max}, x_{min}, and p fixed, the value of $E[u(x)]$ is fixed, so the value of

x_{bond} is also fixed. T   Holding x_{max}, x_{min}, and p fixed, we can therefore use (24) to

examine how $x_{bond}^{ambiguity}$ changes when we change the degree of ambiguity λ and the

degree of ambiguity tolerance α, with x_{bond} remaining fixed.   We find

(25) $\dfrac{d\left[x_{bond}^{ambiguity}\right]}{d\lambda} = \dfrac{[\alpha - p][u(x_{max}) - u(x_{min})]}{u'\left(x_{bond}^{ambiguity}\right)} < 0$

and

(26) $\dfrac{d\left[x_{bond}^{ambiguity}\right]}{d\alpha} = \dfrac{\lambda[u(x_{max}) - u(x_{min})]}{u'\left(x_{bond}^{ambiguity}\right)} > 0$.

The quantity $x_{bond}^{ambiguity} - x_{bond}$ is a measure of the degree to which the equity

premium exceeds that expected from risk aversion alone. Conditions (25) and (26)

indicate that more ambiguity and less ambiguity tolerance each increase the equity

premium. The equity premium increases more significantly when the stake is larger.

5. Conclusion

The distinction between risk and ambiguity has been recognized since Knight (1921) and Keynes (1921). Ellsberg's (1961) seminal experiment demonstrated that the presence of ambiguity can systematically impact decision behavior. However, ambiguity has not been used to explain the puzzling behaviors of decision making under uncertainty as much as it might. This paper has sought to contribute in this regard.

We have illustrated the potential explanatory power of ambiguity. It can explain why the same person may rationally buy insurance and gamble. It can explain why some people will gamble and others will not. It can explain why a gambler may rationally prefer a higher variance bet to a lower variance bet. It can explain why a risk averse decision-maker may rationally exhibit loss aversion.   It can explain why losses may loom larger than gains to a decision maker. It can explain the equity premium puzzle.

In his Noble lecture, Maurice Allais (1990) emphasizes the importance of developing theories with testable implications, and our theory offers a number of these. First, there is more scope for ambiguity to motivate chance taking by risk-averse decision makers when (1) the bad outcome is very likely and the good outcome is not, and (2) the decision maker is more optimistic. More ambiguity or larger stakes will either reinforce risk aversion not taking a chance (when the decision maker is not optimistic enough) or offset the risk aversion and motivate taking a chance (when the decision maker is optimistic enough).   Designing experiments to test and explore these hypotheses will be meaningful future work.

It seems especially challenging, but also important, to distinguish our theory from the theory that utility function convexity (or a preference for risk seeking) explains chance taking. For example, our theory indicates people often prefer betting on single numbers in roulette more than betting on a color because the lower probability p of winning the single number bet requires a lower level of optimism to make taking a chance worthwhile under ambiguity. Utility convexity indicates the single number bet is preferred because the win is more life changing.   Which is correct?

If our theory is empirically validated, it would restore the historically comfortable assumption that the marginal utility of wealth is decreasing over all wealth levels for the typical person.   Alternatively, if further empirical research validates the existence of segments risk seeking behavior in particular contexts, we will better understand those contexts, including investing, gambling, and loss aversion.

6. **References**

Allais, M., 1990, My Conception of Economic Science. Methodus 2, l-5.

Arrow, K. J., & Hurwicz, L. (1972). An optimality criterion for decision-making under ignorance. In C. F. Carter & J. L. Ford (Eds.), *Uncertainty and expectations in economics: Essays in Honour of G.L.S. Shackle*. Oxford: Basil Blackwell.

Barseghyan, L., Molinari, F., O'Donoghue, T., & Teitelbaum, J. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6), 2499-2529.

Benartzi, S. & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics,* 110(1), 75-92.

Dickerson, M. G., Fabre, J. & Bayliss, D. (1986). A comparison of TAB customers and poker machine players. In J. McMillen (Ed.), *Gambling in the 80's.* Griffith University, Queensland: National Association for Gambling Studies.

Dobbs, I. M. (1988). Risk aversion, gambling, and the labour-leisure choice. *Scottish Journal of Political Economy*, 35(2), 171-175.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75(4), 643-669.

Epstein, L., &. Zin, S. (1990). 'First-order' risk aversion and the equity premium puzzle. *Journal of Monetary Economics*, 26(3), 387-407.

Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly Journal of Economics*, 75(4), 670-689.

Flemming, J. S. (1969). The utility of wealth and the utility of windfalls. *Review of Economic Studies,* 36(1), 55-66.

Friedman, M., & Savage, L. J. (1948). Utility analysis of choices involving risk. *Journal of Political Economy,* 56(4), 279-304.

Frisch, D., & Baron, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making*, 1(3), 149-157.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18, 141-153.

Gregory, N. (1980). Relative wealth and risk taking: A short note on the Friedman-Savage utility function. *Journal of Political Economy*, 88(6), 1226-1230.

Hakansson, N. (1970). Friedman-Savage utility functions consistent with risk-aversion. *Quarterly Journal of Economics,* 84(3), 472-487.

Hurwicz, L. (1951). The generalized Bayes minimax principle: A criterion for decision making under uncertainty. Cowles Commission Discussion Paper, Statistics 335.

Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental test of the endowment effect and the Coase theorem. *Journal of Political Economy,* 98(6), 1325-1348.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-291.

Kahneman, D., & Tversky, A. (1984). Choice, values, and frames. *American Psychologist*, 39(4), 341-350.

Katz, E. (1983). Wealth, rank, risk aversion, and the Friedman-Savage utility function. *Australian Economic Papers*, 22(41), 487-490.

Keynes, J. M. (1921). *A treatise on probability*. London, UK: Macmillan.

Knight, F. (1921). *Risk, uncertainty and profit*. New York: Augustus M. Kelley.

Kocherlakota, R. (1996). The equity premium: It's still a puzzle. *Journal of Economic Literature*, 34(1), 42-71.

Koszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4), 1133-1165.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311-328.

Levy, M., & Levy, H. (2001). Testing for risk aversion: A stochastic dominance approach. *Economics Letters,* 71(2), 233-240.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, 60(2), 151-156.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. (PDF). *Journal of Monetary Economics,* 15(2): 145-161.

Melkonyan, T., & Pingle, M. (2010). Ambiguity, pessimism, and rational religious choice.
Theory and Decision, 69(3), 417-438.

Ng, Y. K. (1975). Why do people buy lottery tickets? Choices involving risk and the indivisibility of expenditure. *Journal of Political Economy*, 73(5), 530-535.

Otrok, C., Ravikumar, B., & Whiteman, C. H. (2002). Habit formation: A resolution of the equity premium puzzle? *Journal of Monetary Economics*, 49(6), 1261-1288.

Rabin, M., & Thaler, R. H. (2001). Anomalies: Risk aversion. *Journal of Economic Perspectives*, 15(1), 219-232.

Rietz, T. A. (1988). The equity risk premium: A solution. *Journal of Monetary Economics*, 22(1), 117-131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.

Walker, M. B. (1985). Explanations for gambling. In G. Caldwell, B. Haig, M. G. Dickerson & L. Sylvan (Eds.), *Gambling in Australia* (pp. 146-162). Sydney: Croom Helm.

Walker, M. B. (1992). Irrational thinking among slot machine players. *Journal of Gambling Studies*, 8(3), 245-261.

Appendix

Proof of Proposition 1: With two possible outcome x_{min} and x_{max}, the difference

between the valuation of the minimax expected utility $V(x)$ given by (7) and the expected

utility of the certain outcome $E[u(x)]$ given by (3) is

$$V(x) - E[u(x)] = [1 - \lambda][pu(x_{max}) + [1 - p]u(x_{min})] + \lambda[\alpha u(x_{max}) +$$
$$[1 - \alpha]u(x_{min})] - [pu(x_{max}) + [1 - p]u(x_{min})]$$

Or, equivalently,

$$V(x) - E[u(x)]$$
$$= [pu(x_{max}) + [1 - p]u(x_{min})] - \lambda[pu(x_{max}) + [1 - p]u(x_{min})]$$
$$+ \lambda[\alpha u(x_{max}) + [1 - \alpha]u(x_{min})] - [pu(x_{max}) + [1 - p]u(x_{min})]$$

Which produces

$$V(x) - E[u(x)] = -\lambda[pu(x_{max}) + [1 - p]u(x_{min})] + \lambda[\alpha u(x_{max}) + [1 - \alpha]u(x_{min})]$$

Factoring out λ and $u(x_{max}) - u(x_{min})$ yields

$$V(x) - E[u(x)] = \lambda[\alpha - p][u(x_{max}) - u(x_{min})]$$

Therefore, it must follow that

$$V(x) \left(\genfrac{}{}{0pt}{}{>}{\genfrac{}{}{0pt}{}{=}{<}}\right) E[u(x)] \Leftrightarrow \lambda[\alpha - p][u(x_{max}) - u(x_{min})] \left(\genfrac{}{}{0pt}{}{>}{\genfrac{}{}{0pt}{}{=}{<}}\right) 0$$

Proof Proposition 3

$$V(x) - u(\bar{x}) = [1 - \lambda][pu(x_{max}) + [1 - p]u(x_{min})]$$

$$+ \lambda[\alpha u(x_{max}) + [1 - \alpha]u(x_{min})] - u(px_{max} + [1 - p]x_{min})$$

$$= [pu(x_{max}) + [1 - p]u(x_{min})] - \lambda[pu(x_{max}) + [1 - p]u(x_{min})]$$

$$+ \lambda[\alpha u(x_{max}) + [1 - \alpha]u(x_{min})] - u(px_{max} + [1 - p]x_{min})$$

$$= pu(x_{max}) + u(x_{min}) - pu(x_{min}) - \lambda pu(x_{max}) - \lambda u(x_{min}) + \lambda pu(x_{min})$$

$$+ \lambda \alpha u(x_{max}) + \lambda u(x_{min}) - \lambda \alpha u(x_{min}) - u(px_{max} + [1 - p]x_{min})$$

$$= p[u(x_{max}) - u(x_{min})] - \lambda p[u(x_{max}) - u(x_{min})] + \lambda \alpha[u(x_{max}) - u(x_{min})]$$

$$- u(px_{max} + [1 - p]x_{min}) - u(x_{min})$$

$$= [u(x_{max}) - u(x_{min})][\lambda \alpha + p - \lambda p] - u(px_{max} + [1 - p]x_{min}) - u(x_{min})$$

$$= [u(x_{max}) - u(x_{min})]\left[[\lambda \alpha + [1 - \lambda]p] - \frac{u(px_{max} + [1 - p]x_{min}) - u(x_{min})}{[u(x_{max}) - u(x_{min})]}\right]$$

$$= [u(x_{max}) - u(x_{min})]\left[\lambda[\alpha - p] + \frac{p[u(x_{max}) - u(x_{min})] - u(px_{max} + [1 - p]x_{min}) + u(x_{min})}{[u(x_{max}) - u(x_{min})]}\right]$$

$$= [u(x_{max}) - u(x_{min})]\left[\lambda[\alpha - p] + \frac{pu(x_{max}) + [1 - p]u(x_{min}) - u(px_{max} + [1 - p]x_{min})}{[u(x_{max}) - u(x_{min})]}\right]$$

$$= [u(x_{max}) - u(x_{min})]\left[\lambda[\alpha - p] - \frac{u(px_{max} + [1 - p]x_{min}) - [pu(x_{max}) + [1 - p]u(x_{min})]}{[u(x_{max}) - u(x_{min})]}\right]$$

Since $u(x_{max}) - u(x_{min}) > 0$, it must follow that

$$V(x) \begin{pmatrix} > \\ = \\ < \end{pmatrix} u(\bar{x}) \iff \lambda[\alpha - p] \begin{pmatrix} > \\ = \\ < \end{pmatrix} \frac{u(px_{max}+[1-p]x_{min})-[pu(x_{max})+[1-p]u(x_{min})]}{[u(x_{max})-u(x_{min})]} > 0$$

Proof Proposition 4: If there is some ambiguity (i.e., $\lambda > 0$), and if DM has some ambiguity tolerance (i.e., $\alpha > 0$), then $V(x) = u(\bar{x})$ implies

$$\lambda[\alpha - p] - \frac{u(px_{max} + [1-p]x_{min}) - [pu(x_{max}) + [1-p]u(x_{min})]}{u(x_{max}) - u(x_{min})} = 0$$

Let $f(p)$ denote a differentiable function of the probability p of the good outcome

$$f(p) = \lambda[\alpha - p] - \frac{u(px_{max} + [1-p]x_{min}) - [pu(x_{max}) + [1-p]u(x_{min})]}{u(x_{max}) - u(x_{min})} = 0$$

Then, there exists a unique likelihood $\bar{p}$ of the good outcome such that $p \left(\substack{< \\ = \\ >}\right) \bar{p} \Leftrightarrow$

$$V(x) \left(\substack{> \\ = \\ <}\right) u(\bar{x}).$$

To characterize the shape of $f(p)$ function, we first derive the value of $f(p)$ function evaluated at $p = 0$ and $p = 1$ respectively.

When $p = 0$, then $f(0) = \lambda\alpha - \frac{u(x_{min}) - [u(x_{min})]}{u(x_{max}) - u(x_{min})} = \lambda\alpha > 0$

When $p = 1$, then $f(1) = \lambda[\alpha - 1] - \frac{u(x_{max}) - [u(x_{max})]}{u(x_{max}) - u(x_{min})} = \lambda[\alpha - 1] < 0$

Since $f(p) = \lambda\alpha$ at $p = 0$ is strictly greater than $f(p) = \lambda[\alpha - 1]$ at $p = 1$, the $f(p)$ may possibly be a decreasing function in p. To verify this, we then perform the first derivative $f'(p)$ to measure the rate of change of the function, and the second derivative to measure the rate of change of the first derivative $f'(p)$, respectively.

$$\frac{df(p)}{dp} = f'(\bar{p}) = -\lambda - \left[\frac{u'(px_{max} + [1-p]x_{min})[x_{max} - x_{min}] - u(x_{max}) + u(x_{min})}{u(x_{max}) - u(x_{min})}\right]$$

$$= -\lambda - \frac{u'(px_{max} + [1-p]x_{min})[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})} + \frac{[u(x_{max}) - u(x_{min})]}{u(x_{max}) - u(x_{min})}$$

$$= [1 - \lambda] - \frac{u'(px_{max} + [1-p]x_{min})[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})}$$

$$= [1 - \lambda] - \frac{u'(px_{max} + [1-p]x_{min})}{\frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}}$$

Then, evaluating $f'(\bar{p})$ at $p = 0$ and $p = 1$ yields

$$f'(0) = [1 - \lambda] - \frac{u'(x_{min})}{\frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}} < 0$$

Since $u'(x_{min}) > \frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}$ and $\frac{u'(x_{min})}{\frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}} > 1$, $f'(0)$ then is strictly negative.

$$f'(1) = [1 - \lambda] - \frac{u'(x_{max})}{\frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}}$$

The slope of $f(p)$ evaluate at $p = 1$ indicates $f(p)$ is monotonically increasing in p if

$[1 - \lambda] < \frac{u'(x_{max})}{\frac{u(x_{max}) - u(x_{min})}{[x_{max} - x_{min}]}}$. Then, the second derivative implies

$$\frac{d^2 f(p)}{dp^2} = f''(\bar{p}) = -\frac{u''(px_{max} + [1-p]x_{min})[x_{max} - x_{min}]^2}{u(x_{max}) - u(x_{min})} > 0$$

This assumption of increasing utility $u'(*) > 0$ and risk aversion $u''(*) < 0$ implies the

sign of the second derivative is strictly positive. Therefore, we conclude that the function $f(p)$ is monotonically decreasing at increasing rate as the probability p of the good outcome increases.

The plot of $f(p)$ function against p is provided below.
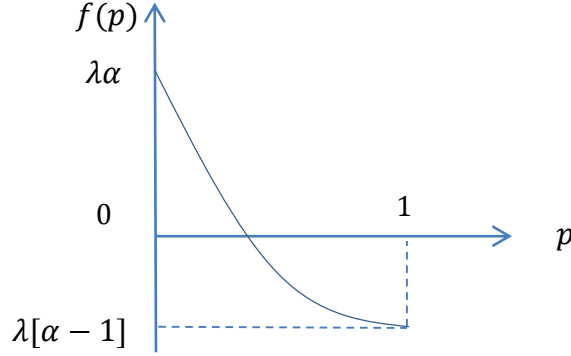
Finally, we then perform comparative statics to examine the impact of changing in ambiguity λ and ambiguity tolerance α on the unique likelihood of $\bar{p}$. Allowing $\bar{p}, \lambda$, and α to change, the total differential of the function $f(p) = 0$ can be written as

$$[\alpha - \bar{p}]d\lambda + \lambda d\alpha - \lambda d\bar{p} - \left[\frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}] - [u(x_{max}) - u(x_{min})]]}{u(x_{max}) - u(x_{min})}\right]d\bar{p} = 0$$

$$[\alpha - \bar{p}]d\lambda + \lambda d\alpha - \lambda d\bar{p} - \left[\frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})} - \frac{[u(x_{max}) - u(x_{min})]}{u(x_{max}) - u(x_{min})}\right]d\bar{p} = 0$$

$$[\alpha - \bar{p}]d\lambda + \lambda d\alpha - \lambda d\bar{p} + \left[-\frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})} + 1\right]d\bar{p} = 0$$

$$\left[[1 - \lambda] - -\frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})}\right]d\bar{p} = -[\alpha - \bar{p}]d\lambda - \lambda d\alpha$$

Dividing both sides of the equation by $1 - \lambda - \frac{u'(\bar{p}x_{max}+[1-\bar{p}]x_{min})[[x_{max}-x_{min}]}{u(x_{max})-u(x_{min})}$ yields

$$d\bar{p} = \frac{-[\alpha - \bar{p}]}{\left[[1-\lambda] - \frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})}\right]} d\lambda$$

$$- \frac{\lambda}{\left[[1-\lambda] - \frac{u'(\bar{p}x_{max} + [1-\bar{p}]x_{min})[[x_{max} - x_{min}]}{u(x_{max}) - u(x_{min})}\right]} d\alpha$$

Knowing the first derivative $f'(\bar{p}) = [1-\lambda] - \frac{u'(px_{max}+[1-p]x_{min})[x_{max}-x_{min}]}{u(x_{max})-u(x_{min})} < 0$, the

multiplier of each variable on the right side must be positive, meaning $\bar{p}$ is increasing in

ambiguity $\left(\frac{d\bar{p}}{d\lambda} > 0\right)$ and in ambiguity tolerance $\left(\frac{d\bar{p}}{d\alpha} > 0\right)$.

To illustrate how the three parameters p, λ, and α affect the willingness of DM to

take the chance, we plot the two functional terms within the $f(p)$ function as they depend

upon the probability p of the good outcome.

$$f(p) = \lambda[\alpha - p] - \frac{u(px_{max} + [1-p]x_{min}) - [pu(x_{max}) + [1-p]u(x_{min})]}{u(x_{max}) - u(x_{min})}$$

The linear term $\lambda[\alpha - p]$ implies the intercept $\lambda\alpha$ when $p = 0$ and the slope of $-\lambda$ when

$0 < p \leq 1$. The curvature term $\frac{u(px_{max}+[1-p]x_{min})-[pu(x_{max})+[1-p]u(x_{min})]}{u(x_{max})-u(x_{min})}$ implies:

When $p = 0$, then $\frac{u(px_{max}+[1-p]x_{min})-[pu(x_{max})+[1-p]u(x_{min})]}{u(x_{max})-u(x_{min})} = \frac{u(x_{min})-[u(x_{min})]}{[u(x_{max})-u(x_{min})]} = 0$

When $p = 1$, then $\frac{u(px_{max}+[1-p]x_{min})-[pu(x_{max})+[1-p]u(x_{min})]}{u(x_{max})-u(x_{min})} = \frac{u(x_{max})-[u(x_{max})]}{[u(x_{max})-u(x_{min})]} = 0$

As the curvature function takes the value of zero when $p = 0$ and when $p = 1$,

the function may have the maximum point. Hence, we perform the first-order derivative

with respect to p and set it up equal zero to identify the condition, where the function will reach the maximum.

$$\frac{d}{dp}\left[\frac{u(px_{max} + [1-p]x_{min}) - [pu(x_{max}) + [1-p]u(x_{min})]}{u(x_{max}) - u(x_{min})}\right] = 0$$

$$\frac{u'(px_{max} + [1-p]x_{min})[[x_{max} - x_{min}] - [u(x_{max}) - u(x_{min})]}{[u(x_{max}) - u(x_{min})]} = 0$$

$$\frac{u'(px_{max} + [1-p]x_{min})[x_{max} - x_{min}]}{[u(x_{max}) - u(x_{min})]} - 1 = 0$$

$$\frac{u'(px_{max} + [1-p]x_{min})[x_{max} - x_{min}]}{[u(x_{max}) - u(x_{min})]} = 1$$

$$u'(px_{max} + [1-p]x_{min}) = \frac{[u(x_{max}) - u(x_{min})]}{[x_{max} - x_{min}]}$$

This condition implies there must exist the maximum at the point where the slope of the utility of the expected value $u(\bar{x})$ equal the slope of the expected utility $E[u(x)]$. This condition is shown in the diagram below.

$u'(\hat{p}x_{max} + [1-\hat{p}]x_{min})$

$[u(x_{max}) - u(x_{min})]$

$x_{max} - x_{min}$
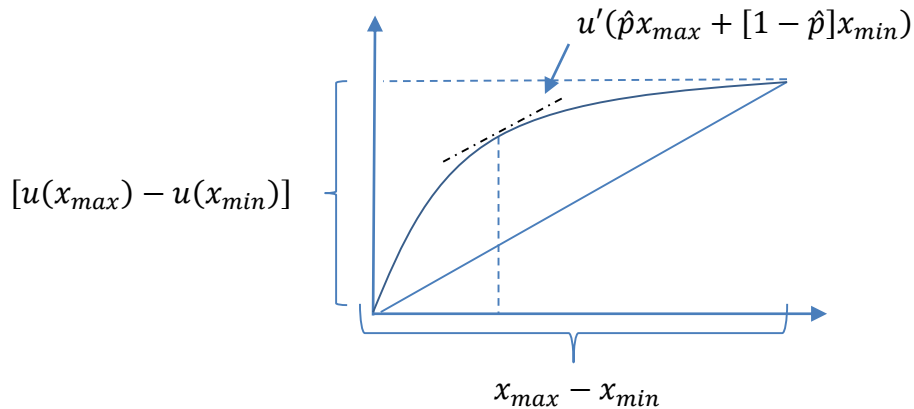
Finally, the plot of the two functional terms within $f(p)$ in condition (9) can be shown as the following diagram.
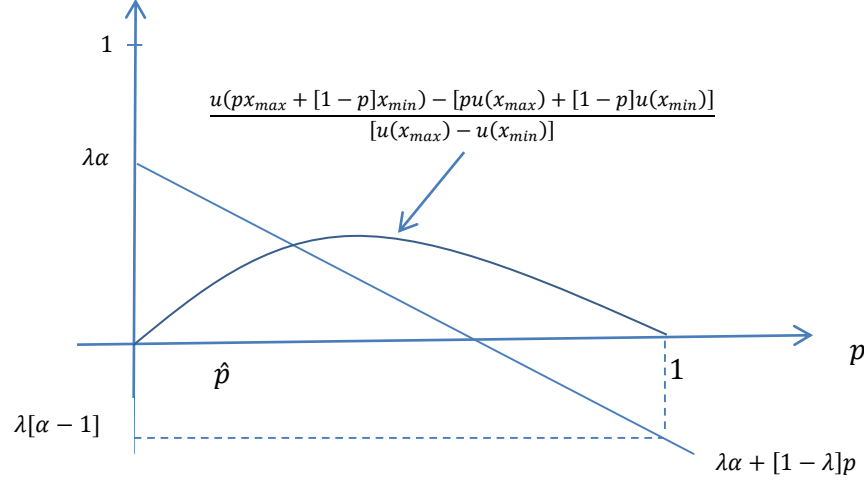
Proof of an Application of Fair Lottery with N participants: To examine whether the minimax expected utility V is greater than the utility of the expected value, we construct $V - u(\bar{x}) > 0$, which implies

$$[1 - \lambda]\left[\left[\frac{1}{N}\right]u(N - 1) + \left[\frac{N - 1}{N}\right]u(-1)\right] + \lambda\big[\alpha[u(N - 1)] + [1 - \alpha][u(-1)]\big]$$

$$> u\left(\frac{1}{N}[N - 1] + \frac{N - 1}{N}[-1]\right)$$

This may be written as

$$\left[\left[\frac{1}{N}\right]u(N - 1) + \left[\frac{N - 1}{N}\right]u(-1)\right] - \lambda\left[\left[\frac{1}{N}\right]u(N - 1) + \left[\frac{N - 1}{N}\right]u(-1)\right]$$

$$+ \lambda\big[\alpha[u(N - 1)] + [1 - \alpha][u(-1)]\big] > u\left(\frac{1}{N}[N - 1] + \frac{N - 1}{N}[-1]\right)$$

Recall $\frac{N-1}{N} = 1 - \frac{1}{N}$, we then can rewrite the equation as

$$\left[\frac{1}{N}\right]u(N-1) + u(-1) - \left[\frac{1}{N}\right]u(-1) - \lambda\left[\frac{1}{N}\right]u(N-1) - \lambda u(-1) + \lambda\left[\frac{1}{N}\right]u(-1)$$

$$+ \lambda\alpha u(N-1) + \lambda[u(-1)] - \lambda\alpha[u(-1)] > u\left(\frac{1}{N}[N-1] + \frac{N-1}{N}[-1]\right)$$

Subtracting both sides by $u(-1)$ and rearranging the equation yields

$$\frac{1}{N}[u(N-1) - u(-1)] - \lambda\left[\frac{1}{N}\right][u(N-1) - u(-1)] + \lambda\alpha[u(N-1) - u(-1)]$$

$$> u\left(\frac{1}{N}[N-1] + \frac{N-1}{N}[-1]\right) - u(-1)$$

Factoring out $[u(N-1) - u(-1)]$ on the left side yields

$$[u(N-1) - u(-1)]\left[\lambda\alpha + [1-\lambda]\frac{1}{N}\right] > u\left(\frac{1}{N}[N-1] + \frac{N-1}{N}[-1]\right) - u(-1)$$

Dividing both sides of the equation by $[u(N-1) - u(-1)]$ yields

$$\lambda\alpha + [1-\lambda]\frac{1}{N} > \frac{u\left(\frac{1}{N}[N-1] + \frac{N-1}{N}[-1]\right) - u(-1)}{u(N-1) - u(-1)}$$

Recall $\frac{1}{N}[N-1] + \frac{N-1}{N}[-1] = 0$, we get

$$\lambda\alpha + [1-\lambda]\frac{1}{N} > \frac{u(0) - u(-1)}{u(N-1) - u(-1)}$$

Proof of an Application of a Roulette Wheel: For $1 bet on red, the expected value of the

bet is

$$\bar{x}_R = \frac{18}{38}[+1] + \frac{20}{38}[-1] = -\frac{2}{38} = -0.0526$$

To examine whether the minimax expected utility is greater than the utility of not playing

the roulette, we then construct $V - u(0) > 0$:

$$(1 - \lambda)\left[\frac{18}{38}u(+1) + \frac{20}{38}u(-1)\right] + \lambda[\alpha u(+1) + (1 - \alpha)u(-1)] > u(0)$$

Recall $\frac{20}{38} = 1 - \frac{18}{38}$, this equation may be written as

$$\left[\frac{18}{38}u(+1) + \left[1 - \frac{18}{38}\right]u(-1)\right] - \lambda\left[\frac{18}{38}u(+1) + \left[1 - \frac{18}{38}\right]u(-1)\right]$$

$$+ \lambda[\alpha u(+1) + (1 - \alpha)u(-1)] > u(0)$$

Or, equivalently,

$$\frac{18}{38}u(+1) + u(-1) - \frac{18}{38}u(-1) - \lambda\frac{18}{38}u(+1) - \lambda u(-1) + \lambda\frac{18}{38}u(-1) + \lambda\alpha u(+1)$$

$$+ \lambda u(-1) - \lambda\alpha u(-1) > u(0)$$

Subtracting both sides by $u(-1)$ and rearranging the equation yields

$$\frac{18}{38}[u(+1) - u(-1)] - \lambda\frac{18}{38}[u(+1) - u(-1)] + \lambda\alpha[u(+1) - u(-1)]$$

$$> u(0) - u(-1)$$

Factoring out $u(+1) - u(-1)$ yields

$$[u(+1) - u(-1)] \left[\lambda \alpha + [1 - \lambda] \frac{18}{38} \right] > u(0) - u(-1)$$

Dividing both sides of the equation by $u(+1) - u(-1)$ yields

$$\lambda \alpha + [1 - \lambda] \frac{18}{38} = \frac{u(0) - u(-1)}{u(+1) - u(-1)}$$

Given the expected value of \$1 bet on a single number $\bar{x}_3 = \frac{18}{38}[+1] + \frac{20}{38}[-1] = -\frac{2}{38} = -0.0526$, we can proceed in a similar manner to derive the condition $V - u(0)$. This condition is given by

$$\lambda \alpha + [1 - \lambda] \frac{1}{38} > \frac{u(0) - u(-1)}{u(+35) - u(-1)}$$

Proof of an Application of Buying Insurance

Given the expected value of the loss is $\bar{x} = p[0] + [1 - p][-a] = -a[1 - p] < 0$, we then construct $V(x) - u(\bar{x})$, which implies

$$(1 - \lambda)[pu(0) + [1 - p]u(-\alpha)] + \lambda[\alpha u(+1) + (1 - \alpha)u(-a)] > u(-a[1 - p])$$

This condition can be written as

$$pu(0) + u(-\alpha) - pu(-\alpha) - \lambda pu(0) - \lambda u(-\alpha) + \lambda pu(-\alpha) + \lambda \alpha u(+1) + \lambda u(-a)$$
$$- \lambda \alpha u(-a) > u(-a[1 - p])$$

Subtracting both sides by $u(-\alpha)$ and rearranging the left side of the equation yields

$$p[u(0) - u(-\alpha)] - \lambda p[u(0) - u(-\alpha)] + \lambda \alpha[u(0) - u(-\alpha)]$$
$$> u(-a[1 - p]) - u(-\alpha)$$

Factoring out $u(0) - u(-\alpha)$ on the left side yields

$$[u(0) - u(-\alpha)][\lambda\alpha + [1 - \lambda]p] = u(-a[1 - p]) - u(-\alpha)$$

Dividing both sides of the equation yields

$$\lambda\alpha + [1 - \lambda]p = \frac{u(-a[1 - p]) - u(-\alpha)}{u(0) - u(-\alpha)}$$

Proof of an Application for the Equity Premium Puzzle

Given the expected return from the stock is $\bar{x} = p\, x_{max} + [1 - p](0)$ and the expected utility

of the stock $E[u(\bar{x})]$ is equivalent to the utility of the expected return from holding the bond

$u(x_{bond})$, we then construct $V(x) - E[u(\bar{x})] = 0$:

$$(1 - \lambda)[pu(x_{max}) + [1 - p]u(0)] + \lambda[\alpha u(x_{max}) + (1 - \alpha)u(0)]$$

$$= p\, u(x_{max}) + [1 - p]u(0)$$

This can be written as

$$pu(x_{max}) + [1 - p]u(0) - \lambda[pu(x_{max}) + [1 - p]u(0)] + \lambda[\alpha u(x_{max}) + (1 - \alpha)u(0)]$$

$$= p\, u(x_{max}) + [1 - p]u(0)$$

Subtracting both side of the equation by $pu(x_{max}) + [1 - p]u(0)$ yields

$$-\lambda[pu(x_{max}) + [1 - p]u(0)] + \lambda[\alpha u(x_{max}) + (1 - \alpha)u(0)] = 0$$

This condition can also be written as

$$-\lambda pu(x_{max}) - \lambda u(0) + \lambda pu(0) + \lambda\alpha u(x_{max}) + \lambda u(0) - \lambda\alpha u(0) = 0$$

Rearranging this equation, we have

$$-\lambda p[u(x_{max}) - u(0)] + \lambda \alpha[u(x_{max}) - u(0)] = 0$$

Factoring out $u(x_{max}) - u(0)$ and λ on the left side yields

$$[u(x_{max}) - u(0)]\lambda[\alpha - p] = 0$$

Therefore, $V(x) \begin{pmatrix} > \\ = \\ < \end{pmatrix} E[u(x)]$ if and only if $[u(x_{max}) - u(0)]\lambda[\alpha - p] \begin{pmatrix} > \\ = \\ < \end{pmatrix} 0$

Chapter 2

The Effects of Communication and Punishment on Trust and Trustworthiness

Abstract

The paper experimentally examines the effects of punishment, communication and the interaction of these on trust and trustworthiness. Consistent with previous works, trusting and trustworthy behaviors are observed across all different treatments, and subjects are not motivated solely by self-interest. The presence of either the punishment option or the communication option crowds out trustworthiness which trust can elicit, or reduces the marginal effect of trust. However, if trustor expresses the highest level of trust accompanied by communication without implementing any punishment threat, the trustee's trustworthiness significantly increases.

1. Introduction

Concerns about trust have grown substantially during the last two decades, as the development of trust helps people to establish productive and meaningful relationships. In business and economics, "trust always affects two outcomes: speed and cost. When trust goes down, speed goes down and cost goes up" (Covey, 2006, p.22). There is a great deal of supporting evidence that trust fosters business and economic activities by lowering monitoring costs (e.g. Frank, 1988), lowering turnover (Dirks & Ferrin, 2002), enhancing economic outcomes (Fukuyama, 1995; Putnam, 1993), and improving the overall economy (Knack & Keefer, 1997). Not surprisingly, such observations have led many economists to examine the determinants of trust, especially in the experimental setting.

Beginning with Berg et al.'s (1995) study of trust and reciprocity, the trust game has become a popular tool in measuring trust and trustworthiness in both controlled laboratories and filed experiments (e.g., Barr, 2003; Glaeser et al., 2000; Cochard, Nguyen and Willinger, 2004 Willinger, Keser, Lohmann and Usunier, 2003; Coricelli, Morales and Mahlstedt, 2006; Bacharach, Guerra and Zizzo, 2007; Bellemare and Kroger, 2007; Bigoni, Bortolotti, Casari, Gambetta and Pancotto, 2013). Recently, a body of economic literature on this topic has focused on understanding the extent to which punishment and communication affect trust and trustworthiness (e.g., Fehr and Gächter, 2000; Bohnet et al., 2001: Fehr and List 2004; Charness and Dufwenberg, 2006; Kimbrough, et al., 2007; Schotter & Sopher, 2007).

The ability to punish and the ability to communicate may alter players' beliefs and actions. The effectiveness of each mechanism is also dependent upon several factors such as the cost of punishment (Rigdon, 2009), the number of participants (Bohnet et al., 2001; Charness et al., 2008), and the form of communication (Issac and Walker, 1998; Duffy and Feltovich, 2002). However, these studies have examined either punishment or communication, but not both, meaning one must speculate about how they interact. To date, no trust game experiment has offered a comprehensive design, in which subjects are allowed to have both the ability to punish and the ability to communicate.

Another aspect of the trust game experiment that has been absent in the literature is the possibility of having one-way communication beginning from the trustor. Previous experiments have examined one-way communication beginning from trustee (e.g., Bracht and Feltovich, 2009), and two-way communication where subjects have the opportunity to dialogue before making decisions (e.g., Issac and Walker, 1998; Duffy and Feltovich 2002).

This study makes two important contributions. First, the study sheds light on how punishment and communication interact in terms of impacting trust and trustworthiness. Subjects in the role of trustor are randomly assigned into one of the four conditions: punish only, communicate only, punish or communicate, and neither. By looking at punishment and communication separately and jointly, we can identify any interaction effect. Second, this study broadens the literature and offers a new experimental design that allows one-way communication in the form of written message by trustor.

The remainder of this study is structured as follows. Section 2 specifies our experimental design, our subject pool, and the details of the experimental parameters and procedures. Section 3 provides the predictions of self-interest, related evidence from previous studies, and tests of the self-interest hypothesis using data from our study. Section 4 presents and analyzes the data from our four experimental treatments, focusing on understanding the effects of punishment and communication. Section 5 summarizes and concludes the paper.

2. Experimental Design and Procedures

The experiment contains four treatments and uses a 2 x 2 design, as presented in Table 1. The NPNC treatment is the standard trust game, which offers neither the ability to punish nor the ability to communicate. NPNC provides a baseline measure of trust and trustworthiness, a control against which the following three treatments can be compared. In the WPNC treatment, Actor 1 may punish Actor 2 in the trust game for not providing the desired back transfer, but Actor 1 does not have the ability to communicate a written message to Actor 2. Conversely, in the NPWC treatment, Actor 1 may communicate a written message to Actor 2 but cannot punish Actor 2. The WPNC and NPWC treatments provide baselines against which the results of the WPWC treatment can be compared. In the WPWC treatment, Actor 1 has both the ability to communicate a written message and the ability to punish.

Table1: Experimental Treatments

	No Ability to Communicate	Ability to Communicate
No Ability to Punish	NPNC	NPWC
Ability to Punish	WPNC	WPWC

By comparing the results of the WPWC treatment to those of the other three treatments, we can examine how punishment and communication interact. It could be that they complement each other, so that having both options facilitates trust, or trustworthiness, or both. However, many other outcomes are possible. For example, it could be that taking advantage of the ability to punish cancels out a positive effect of an ability to communicate. The treatments are now described in more detail.

In the NPNC treatment, Actor 1 and Actor 2 are paired and play a standard trust game. Paired participants are in different rooms, and understand their identity is never revealed to the other. Actor 1 and Actor 2 each receive an endowment of 10 shanks (experimental currency units). By writing on a decision sheet, Actor 1 choses a transfer amount $x \in \{0, 1, 2, \ldots, 10\}$ shanks to send to Actor 2 and indicates a desired back-transfer $\hat{y} \in \{0, 1, 2, \ldots, 3x\}$. The decision sheet is then taken to the other room and given to Actor 2. Actor 2 receives 3 shanks for each shank transferred by Actor 1. Then, Actor 2 choses the actual back-transfer level $y \in \{0, 1, 2, \ldots, 3x\}$. The payoff received by Actor 1 is given by $\pi_1 = 10 - x + y$, whereas the payoff for Actor 2 is $\pi_2 = 10 + 3x - y$. Each participant was paid two U.S. dollars for each shank earned (See Appendix A).

The WPNC treatment is identical to the NPNC treatment, except Actor 1 in the WPNC treatment can implement a "conditional payoff cut." The conditional payoff cut is a fixed penalty of 4 shanks $(f = 4)$ imposed on Actor 2 whenever the actual back transfer paid by Actor 2 is less than the desired back transfer of Actor 1 (i.e., whenever $y < \hat{y}$). Importantly, Actor 1 is not required to implement the punishment, and the 4 shanks penalty is not given to the Actor 1 when Actor 1 does not implement the punishment. Actor 2 is aware that Actor 1 may or may not impose the conditional payoff cut. Prior to making the back transfer choice, Actor 2 is informed on the decision sheet whether or not Actor 1 has chosen to implement this conditional payoff cut. As in the NPNC treatment, the Actor 1's payoff is given by $\pi_1 = 10 - x + y$. For Actor 2, the payoff is $\pi_2 = 10 + 3x - y$ as in the NPNC treatment if either Actor 1 does not impose the conditional payoff cut or Actor 2 provides a back transfer that is greater than or equal to the desired back transfer of Actor 1 (i.e., if $y \geq \hat{y}$). Alternatively, the payoff for Actor 2 is $\pi_2 = 10 + 3x - y - 4$ if Actor 1 imposes the conditional payoff cut and Actor 2 fails to provide the actual back transfer that is greater than or equal to the desired back-transfer requested by Actor 1 (i.e., if $y < \hat{y}$).

The NPWC is identical to the NPNC treatment except Actor 1 could provide a written message to Actor 2 prior to submitting the transfer choice. Importantly, Actor 2 is not allowed to provide a message back to Actor 1, so the communication is one-way. Actor 1 completes a decision sheet as in the NPNC control treatment, but the decision sheet in the NPWC treatment has a box indicating either "yes" if Actor1 chooses to communicate by a message or "no" if Actor 1 chooses not to do so. If Actor 1 chooses to

communicate, the message is typed into a computer and sent through a computer network to the computer in front of Actor 2 in the other room. Once Actor 2 receives the decision sheet, if the decision sheet indicates there is a message, then Actor 2 must view the message sent through the computer network prior to deciding upon the actual back transfer. Regardless of whether a message is sent, the Actor 1's payoff is given by $\pi_1 = 10 - x + y$ and the Actor 2's payoff is given by $\pi_2 = 10 + 3x - y$. That is, the only difference between the control treatment NPNC and the communication treatment NPWC is Actor 1 may communicate with Actor 2 by a written message.

The treatment WPWC combines the WPNC treatment with the NPWC treatment. Actor 1 has the ability to punish Actor 2 for not satisfying his desired back-transfer, and Actor 1 can communicate a message to Actor 2. Similar to other treatments, the payoff to Actor 1 is given by $\pi_1 = 10 - x + y$. Since communication through a message sent by Actor 1 has no effect on Actor 2's payoff, the payoff for Actor 2 in this treatment is determined exclusively by whether or not the conditional payoff cut is imposed. Hence, the Actor 2's payoff is $\pi_2 = 10 + 3x - y - 4$ if the Actor1 imposes the conditional payoff cut and if Actor 2 choses back transfer $y < \hat{y}$, while the payoff to Actor 2 is $\pi_2 = 10 + 3x - y$ if either Actor 1 chooses not to impose the conditional payoff cut or if Actor 1 imposes the punishment but Actor 2 choses back transfer $y \geq \hat{y}$.

The experimental data was collected in a series of 10 sessions, where each session randomly assign subjects into different treatments. A total of 260 subjects (130 pairs) participated in the four treatments including 50 subjects (25 pairs) in treatment NPNC, 78 subjects (39 pairs) in treatment WPNC, 52 subjects (26 pairs) in treatment NPWC, and 80

subjects (40 pairs) in treatment WPWC. This included undergraduate students, graduate students, and employees at University of Nevada, Reno. Upon agreeing to participate in the study, subjects completed a demographic questionnaire mainly for the purpose of controlling for individual heterogeneity in our econometric analysis (See Appendix B).

In addition, specific efforts were taken to recruit employees and students with supervisory experience, so the behavior of those with supervisory experience could also be compared to the behavior of those without supervisory experience. Defining a supervisory position as one in which the work of at least one other employee was supervised, there were 128 subjects in supervisory position and 132 subjects in non-supervisory position. However, our results obtained from a variety of statistical tests (e.g., the one-tailed t-test, two-tailed t-test, and regression analysis) indicate no significant effects of supervisory experience and demographic factors (e.g., age, gender, income, education level, religious) on trusting and trustworthy behavior.

3. Underlying Theory and Previous Research

3.1 Null Hypotheses under Pure Self Interest

The degree of trust in the experiment is captured by the willingness of Actor 1 to transfer a positive amount $(0 < x \leq 10)$ of initial endowment to Actor 2, and the degree of trustworthiness is the willingness of Actor 2 to back transfer some amount $(0 < y \leq 3x)$ of the tripled transfer. An important theory of how people will behave is offered by the Nash equilibrium solution concept, which assumes each player is self-interested and acts under the assumption all other players are also self-interested. The following presents a series of hypotheses based upon this theory of pure self-interest.

H1: *Actor 1 in the NPNC treatment will demonstrate no trust by choosing a transfer level of zero.*

If Actor 2 is purely self-interested, then Actor 2 will not provide any back transfer. Anticipating this, Actor 1 will not transfer any value to Actor 2. So, the Nash Equilibrium prediction is $x = 0, y = 0$. The communication of Actor 2 in the form of the chosen value of $\hat{y}$ is cheap talk, for it has no bearing on the outcome of Actor 2, so the theory provides no prediction for the value of $\hat{y}$.

H2: *The ability to impose a conditional payoff cut in the WPNC and WPWC treatments will be implemented, and the back transfer will be either 3 or 4 shanks.*

In the WPNC and WPWC treatments, Actor 1 has the opportunity to penalize Actor 2 by imposing the conditional payoff cut of 4 shanks whenever the actual back transfer is less than the desired back transfer. If both Actor 1 and Actor 2 are self-interested, and if Actor 1 anticipates Actor 2's self-interest, it is obvious that implementing the payoff cut is a dominant strategy for Actor 1, because it guarantees a higher payoff ($f = 4 > f = 0$). Knowing Actor 1's dominant strategy, Actor 2's best response can be either sending back the requested amount when the desired back transfer is less than or equal to 4 shanks, or sending back nothing when the desired back transfer is greater than 4 shanks. With the ability to recognize Actor 2's best response, Actor 1 should appropriately select a combination of the transfer and desired back transfer such that his payoff in the game is maximized. That is, Actor 1 will transfer $x = 1$ and request $\hat{y} = 3$, or transfer $x = 2$ and request $\hat{y} = 4$. Subsequently, there exists two equilibria in the game with punishment, including ($x = 1, \hat{y} = 3, y = 3$) and ($x = 2, \hat{y} = 4, y = 4$).

H3: *In the NPWC and WPWC treatments, the opportunity of Actor 1 to communicate a message to Actor 2 prior to submitting a transfer choice will neither affect the transfer (trust) of Actor 1 nor the back transfer (trustworthiness) or Actor 2.*

The opportunity to communicate through a message is cheap talk. Because it cannot impact Actor 2 in terms of the payoffs, it will not impact the decision of Actor 2. Thus, in the NPWC treatment, when no punishment option is available to Actor 1, Actor 2 will chose a zero back transfer and Actor 1 will choose a zero transfer level. That is, the predictions for the NPWC treatment are the same as that for the NPNC treatment.

H4: *When punishment and communication interact, subjects will not exhibit trust and trustworthiness differently from the case when the punishment option is available separately.*

The predictions for the WPWC treatment are the same as for the WPNC treatment, since the communication is cheap talk.

3.2 Testing the Pure Self-Interest Hypotheses

From previous research using the trust game, we expect the pure self-interest hypotheses to be rejected. Table 2 summarizes the predictions of pure self-interest hypotheses, and Table 3 provides the mean behavior of subjects in our four experimental treatments. As expected, all of the pure self-interest hypotheses are rejected.

In the four treatments, the average transfer of Actor 1 ranges between 60 and 70 percent the 10 shank endowment, and the average back transfer chosen by Actor 2 ranges between 7 and 12 shanks. The difference in the average transfer of Actor 1 across the treatment is not statistically significant, but the average back transfer of Actor 2 does significantly differ ($p < 0.05$).

Table 2: Pure Self-Interest Hypotheses

Treat ment	Transfer x (Shanks)	Punishment (% imposed)	Communication		Actual back transfer y (shanks)
			Desired back transfer $\hat{y}$ (shanks)	Message (% sent)	
NPNC	0	N/A	0	N/A	0
WPN C	1	100%	3	N/A	3
	2	100%	4	N/A	4
NPW C	0	N/A	0	Indifferent	0
WPW C	1	100%	3	Indifferent	3
	2	100%	4	Indifferent	4

Table 3: Experimental Results

Treatme nt	Transfer x (Shanks)	Punishment (% imposed)	Communication		Actual back transfer y (shanks)
			Desired back transfer $\hat{y}$ (shanks)	Message (% sent)	
NPNC $N=24$	7.00	N/A	14.91	N/A	11.95
WPNC $N=37$	6.16	62%	12.67	N/A	7.11
NPWC $N=26$	6.77	N/A	13.80	70%	9.07
WPWC $N=39$	6.69	77%	13.79	85%	7.89

The positive transfer in NPNC treatment (67.2% of the initial endowment) is

consistent with the evidence gathered in previous studies. Berg et al. (1995), who

introduced the trust game, found Actor 1 transferred an average of 52% of the

endowment. Cox (2004) found that 80% of Actor 1 participants (26 out of 32 subjects)

transferred positive amounts and averaged a transfer of 60% of the endowment. Bohet

and Baytelman (2007) reported an average transfer of 55% of the endowment.[1] Our

NPNC treatment is unique in that no previous standard trust game required Actor 1 to

---

[1] See. Camerer (2003), Sapienza, Simats, & Zingales, 2007, Cooper and Kagel (2009), Johnson and Mislin (2011), Eckel and Wilson (2010), and Eckel and Grossman (2008) for reviews of experimental results in the trust game.

provide a requested back transfer.   However, the results in the baseline NPNC treatment are consistent with previous results.

Why do people exhibit significant trust, and not the pure self-interest behavior implicitly associated with the Nash equilibrium prediction?   The bulk of work in experimental economics suggests people exhibit "social preferences," or a concern for the other. One can classify preferences that depart from the self-interest into three categories: positive reciprocity (Ortmann, Fitzgerald & Boeing, 2000; McCabe, Rigdon, & Smith, 2003; Cox, 2004;), inequality aversion or fairness (Rabin, 1993; Fehr & Schmidt, 1999; Bolton & Ockenfels, 1999) and altruism (Cox, Sadiraj, & Sadiraj 2001; Andreoni & Miller, 2002; Charness & Rabin, 2002; Carpenter, Connolly & Myers, 2008).   Our purpose here is not to test one of these theories against the others, but rather is to obtain insight as to how communications and punishment options influence outcomes when social preferences are present.

In Table 3, notice the desired back transfer of Actor 1 is roughly twice the amount sent, or roughly two-thirds the tripled amount received by Actor 2. This implies the average Actor 1 participant is proposing to divide the surplus created by the trust exhibited by Actor 1 in such a manner that the ultimate payoffs of the two players are roughly equal. This is an interesting result in that it helps us recognize Actor 1 can communicate with Actor 2, even when there is no opportunity to send a written message. A larger transfer amount x communicates a greater level of trust, and the desired back transfer level $\hat{y} = 2x$ indicates a desire for fairness in the form of equal outcomes.

Our result for the average desired back transfer is consistent with that obtained from previous studies. Fehr and Rockenbach (2003), in a study of the detrimental effects

of punishment, reported that Actor 1 subjects requested 60 to 67 of the tripled investment as a back transfer. Similarly, Fehr and List (2004) reported the desired back transfer of roughly two-thirds the tripled investment. These have motivated researchers to use the desired back transfer as a means for understanding the motives and preferences of Actor 1, including inequality aversion (Neaf & Schuup, 2009), positive reciprocity (Bicchieri et al., 2010), and guilt aversion (Battigalli & Dufwenberg, 2007).

As shown in Table 3, the average back transfer of Actor 2 in the baseline NPNC treatment is 11.88 shanks. This is significantly higher ($p < 0.05$) than the average back transfers in the other treatments (9.07 in NPWC, 8.10 in WPNC, and 8.03 in WPWC) according to the one-tailed t-test. In contrast to what one might expect, these results suggest the ability to punish and the ability to communicate each hamper the ability to elicit trustworthiness.

Why might the ability to punish a lack of trustworthiness actually reduce trustworthiness? One explanation is positive reciprocity. If people respond to trust by reciprocating with trustworthiness, then trust will elicit trustworthiness. Threating punishment reduces or, as in the case of the Nash Equilibrium described above, entirely eliminates trust from the human interaction, and thereby reduces or eliminates the ability of trust to elicit trustworthiness.

Refraining from implementing punishment when the option is available might be perceived as an additional act of trust. This possibility was examined by Fehr and Rockenbach (2003). They found one third of Actor 1 participants (15 out of 45 subjects) in a punishment treatment voluntarily refrained from using the option. In response, all of

the Actor 2 participants paired with these non-punishing Actor 1 subjects provided a positive back transfer. Fehr and List found the average back transfer was highest (61% of the tripled investment) when the punishment option was available but not used, and they found it was lowest (33% of the tripled investment) when the punishment option was available and used.

The effects of punishment have also been examined in other ways. Rigdon (2009) examined the effect of punishment on the decision of Actor 1. Implementing the punishment threat increased the average transfer of Actor 1, especially when the cost of implementation was relatively cheap. Bohnet et al. (2001) and Charness et al. (2008) examined the effect of third-party punishment. They found implementing a punishment threat from a third party can increase the transfer (trust) of Actor 1 and back transfer (trustworthiness) of Actor 2.

In summary, the research on punishment indicates the ability to punish may crowd out the positive reciprocity which trust can elicit, and the crowd out effect can be reduced or eliminated if the opportunity to punish is explicitly not implemented. The ability to punish may increase the willingness to trust. Third party punishment may increase both trust and trust worthiness.

Research examining communication in the trust game is abundant (e.g., Charness & Dufwenberg, 2006; Kimbrough, et al., 2008; Schotter and Sopher, 2006; Cason, Sheremeta and Zhang, 2012; Sheremeta and Zhang, 2013). However, analogous to the work that has been done on punishment, it appears no one has yet examined whether having the opportunity to communicate but not using it will affect trustworthiness.

Research in business and management can be used to form a hypothesis. Communication

fosters trust and trustworthiness which allows the gains from cooperative behavior to be

captured (e.g., Allert and Chatterjee, 1997; Massey and Kyriazis, 2007; Kottila and

Ronni, 2008, Webster and Wong, 2008). Thus, we might expect not communicating

when the opportunity is available to discourage trustworthiness, even when trust is

demonstrated.

The effect of communication in the form of written message by Actor 1 has also

not been fully examined. Charness and Dufwenberg (2006) did consider written

messages, and they considered three conditions: no message, message from Actor 1 to

Actor 2, and message from Actor 2 to Actor 1. [2] The message transmitted from Actor 2

to Actor 1 was the most effective in terms of promoting cooperative behavior (trust

combined with trustworthiness), and the message that particularly contained a statement

of "promise" was most effective in this regard.

In contrast to our one way communication experiment, or that of Charness

Dufwenberg, most studies examine two-way communication. For instance, Ben-Ner et

al. (2011) reported, when subjects had an opportunity to dialogue and exchange their

proposals prior to making decisions, the average transfer substantially increased (from

50% to 92% of the endowment), and the back transfer significantly increased (from 42%

---

[2] Unlike the standard trust game, subjects in the role of Actor 1 were allowed to choose either "In" or "Out", and once Actor 1 chose "In", Actor 2 would have an opportunity to choose either "Roll" or "Don't Roll" a six-sided die. Therefore, the ultimate payoffs of both players were dependent upon a chance of success or failure when the responder rolled the six-sided die. When the communication goes from Actor 1 to Actor 2, Actor 1 participants chose "In" 67% and Actor 2 subjects chose "Roll" 34% of the time, whereas Actor 1 participants chose "In" 56% and Actor 2 participants chose "Roll" 44% of the time when the communication is not available. There is no statistically significant difference between the two treatments. In the treatment where the communication goes from Actor 2 to Actor 1, Actor 2 participants significantly increased choosing "In" to 74% and Actor 2 participants increased choosing "Roll" to 67% of the time.

to 59% of the tripled investment).[3] Alternatively, Bracht and Feltovich (2009) allowed a cheap talk treatment, where Actor 2 had an opportunity to send a message to Actor 1, and observed that cheap talk has little positive impact on the frequency of Actor 1's investment (40 – 58% of the time) and Actor 2's back transfer (40 – 48% of the time).[4]

Why is communication not just cheap talk as predicted by the theory of pure self-interest?   Bichierri el al. (2010) proposes communication is an effective mechanism to enhance trust and trustworthiness when it is perceived as a promise with some binding force. Cohen et al. (2010) proposes that effective communication expresses a strong concern for social and interpersonal norms related to fairness. Issac and Walker (1998) and Duffy and Feltovich (2002) propose and provide experimental evidence that communication is more effective when subjects can communicate freely face to face, as compared to the communication that takes place in form of numerical or written message.

In summary, communication in the form of a one-way written message may impact behavior, but it is a relatively weak form of communication.   Even when there is no enforcement mechanism, social norms and past experiences imbedded in the psyche of the receiver may give the communication some force. Communication can provide a connection, and merely connecting may influence behavior. Moreover, different words

---

[3] Ben-Ner et al. conducted an experiment that involved four different conditions of communication: 1) no communication; 2) a one-stage computerized negotiation between Actor 1 and Actor 2; 3) a three stage computerized negotiation; and 4) a computerized pre-play chat.

[4] Bracht and Feltovich also reported that the frequency of Actor 1's investment and Actor 2's back transfer increased substantially to 71.2% and 87.7% of the time respectively in the interaction treatment, where cheap talk and observed previous actions were allowed.

provide the opportunity to connect in different ways, and some ways may elicit more trustworthiness than others.

In our treatments, most Actor 1 subjects used punishment when it was available (62% in WPNC and 77% in WPWC), and most Actor 1 subjects use communication when it was available (70% in NPWC and 85% in WPWC). Notice that adding the opportunity to communicate with a written message (moving from WPNC to WPWC) increased the willingness to use punishment. Also, notice that adding the opportunity to punish (moving from NPWC to WPWC) increased the use of written communication. That is, punishment and communication opportunities complemented each other in terms of the prevalence of their use. Yet, examining the average back transfers, we find that the increased use of punishment and communication in the move to the WPWC treatment was not particularly fruitful. Trustworthiness was most effectively elicited when neither punishment nor the ability to communicate through written message was available.

The next section examines the experimental results of the four treatments in more detail, seeking to extract understanding of how punishment and communication, in the specific forms presented here, impact trust and trustworthiness.

4. Analysis of Experimental Results

4.1 Descriptive Statistics

In order to shed the light on the issue of how the use of punishment and communication impacts trust and trustworthiness, we begin our analysis by separating the data of each experimental treatment according to whether Actor 1 adopts the ability to punish, the ability to communicate, or the combined capability. In Table 4, we present a

comparison of subjects' behavior according to the following three main variables: i) the transfer sent by Actor 1; ii) the back transfer chosen by Actor 2; and iii) the return to trust, or the ratio of the back transfer to the transfer.

Table 4 indicates the presence of trust and trustworthiness in every experimental setting. The average transfer of Actor 1 ranges between 50 and 70 percent of the 10 shank endowment, and the average back transfer chosen by Actor 2 ranges between 7 and 13 shanks. The average return to trust ranges from 80 to 200 percent of the transfer. However, the back transfer tends to exhibit higher volatility than other variables according to the estimated coefficient of variation. The high volatility of the back transfer is likely if the Actor 1 subject does not communicate when communication is possible (1.7 in NPWC and 1.3 in WPWC).

Irrespective of whether the ability to punish or the ability to communicate is used, the average transfer of Actor 1 is not significantly different from the baseline NPNC treatment at any conventional level according to the one-tailed t-test. Thus, we conclude Actor 1 subjects who adopt either the ability to punish or the ability to communicate do not behave differently from the condition, where neither of the two mechanisms is available.

Looking at the transfer variable we also learn that the degree of trust is positively related to the use of the communication option but inversely related to the use of the punishment option. As shown in Table 4, the average transfer is significantly lower (6.8 to 5.8 shanks, $p < 0.01$) in the WPNC treatment if the ability to punish is used. Conversely, the average transfer is significantly higher (5.6 to 7.3 shanks $p < 0.05$) in the WPNC treatment if the ability to communicate is used.

Table 4: Splitting Out Communication and Punishment

Variables	NPNC	WPNC			NPWC			WPWC				
	All Subjects	*All Subjects*	Not Use Punishment	Use Punishment	*All Subjects*	Not Use Communication	Use Communication	*All Subjects*	Use Punishment Not Communication	Use Communication Not Punishment	Use Both	Use Neither
Transfer (x)												
Mean	7.0	6.2	6.8	5.8	6.8	5.6	7.3	6.7	5.0	8.6	6.4	7.5
S.D.	3.5	3.3	3.6	3.1	3.6	4.7	3.1	3.0	3.7	2.4	2.9	3.5
C.V.	0.5	0.5	0.5	0.5	0.5	0.8	0.4	0.5	0.7	0.3	0.5	0.5
Back transfer (y)												
Mean	12.0	7.1	7.3	7.0	9.1	4.8	11.0	7.9	7.0	11.4	6.7	12.5
S.D.	7.6	5.6	6.3	5.3	8.3	7.9	8.0	6.3	8.9	6.3	5.8	3.5
C.V.	0.6	0.8	0.9	0.8	0.9	1.7	0.7	0.8	1.3	0.5	0.9	0.3
Return to trust (y/x)												
Mean	1.7	1.4	1.1	1.5	1.3	0.8	1.5	1.2	1.5	1.3	1.1	2.0
S.D.	0.6	0.9	0.6	1.0	0.8	0.9	0.7	0.9	1.3	0.7	0.8	1.4
C.V.	0.3	0.7	0.6	0.7	0.6	1.1	0.5	0.7	0.9	0.5	0.8	0.7
Observations	24	37	14	23	26	8	18	39	4	7	26	2

Notes: Amounts shown are in shanks. S.D. stands for standard deviation. C.V. stands for coefficient of variation (S.D./Mean).

Interestingly, this observed behavior arises in the WPWC treatment. Actor 1 subjects who communicate but refrain from using the punishment transfer the high average of 8.6 shanks, while Actor 1 subjects that do not communicate but adopt the punishment transfer the low average of 5.0 shanks. This difference in the transfer is statistically significant ($p < 0.01$) according to the one-tailed t-test.

Consider the back transfer chosen by Actor 2. The average back transfer is highest when the ability to punish and the ability to communicate are not available, or available but not used (12.0 in NPNC and 12.5 in WPWC). Correspondingly, the average return to trust is also highest when these two mechanisms are not available, or available but not used (1.7 in NPWC and 2.0 in WPWC). This result provides an indication that the presence of both punishment and communication options may place a pressure on Actor 2 to exhibit less trustworthiness.

Knowing the availability of punishment or communication negatively impact trustworthy behavior on average, it is of interest to further explore whether the degree of trustworthiness depends upon whether the available option is adopted by Actor 1. In particular, our results suggest the use of the communication option is an effective mechanism in eliciting Actor 2's trustworthiness while the use of the punishment option is not.

In the NPWC treatment, the average back transfer of Actor 2 is 11.0 shanks if Actor 2 receives a message from Actor 1, while it is only 4.8 shanks if Actor 1 stays silent. This difference is statistically significant at $p < 0.01$ according the one-tailed t-test. When facing punishment in the NPWC treatment, Actor 2 subjects transfer back an average of 7.0 shanks, while the average back transfer is 7.3 shanks when Actor 1

chooses not to apply the punishment. This difference is, however, not statistically

significant according to the one-tailed t-test. Similarly, in the WPWC treatment, the

average back transfer is 11.4 shanks if Actor 2 receives a message but does not face the

punishment, while it is only 7.0 shanks if Actor 1 does not send the message but impose

the punishment threat. This difference is statistically significant at $p < 0.10$ according to

the one-tailed t-test.

4.2 Regression Analysis

In this section, we report the results of several ordinary least-squares regressions

using the entire dataset from the four experimental treatments. This gives us the

opportunity not only to increase the statistical power of hypothesis tests, but also to

separately examine the availability and usage effects of punishment and communication

on the relationship between trust and trustworthiness.

Our regression model specifications (A) and (B) examine the treatment effects on

Actor 2's trustworthiness, or the effects of being in the treatments where the punishment

and communication options are available. The dependent variable is the back transfer of

Actor 2 y_i, which is regressed on a set of dummy variables $T_{k,i}$ indicating the treatment

status, so

(1) $\quad y_i = \sum_{k=1}^{4} \beta_k T_{k,i} + \varepsilon_i.$

This model specification, however, does not allows for the possibility that there may be

an interaction that occurs between the availability of the option and the degree of trust

exhibited by Actor 1. To capture the interaction effect, our regression model

specifications (C), (D) and (E) include another set of interaction terms between the treatment dummy $T_{k,i}$ and the transfer x_i sent by Actor 1.

$$(2) \quad y_i = \sum_{k=1}^{4} \beta_k T_{k,i} + \sum_{k=1}^{4} \gamma_k T_{k,i} x_i + \varepsilon_i.$$

Note that the coefficient γ_k on the interaction term measures the marginal effect of trust on trustworthiness in each treatment (i.e., the change in the back transfer y with respect to a one unit change in the transfer x), or the degree of positive reciprocity exhibited by Actor 2.

Finally, our regression model specifications (F), (G), and (H) examine the usage effect of each mechanism on Actor 2's trustworthiness by introducing three additional interaction variables that indicate whether the ability to punish, the ability to communicate, or the combined capability is implemented.

$$(3) \quad y_i = \sum_{k=1}^{4} \beta_k T_{k,i} + \sum_{k=1}^{4} \gamma_k T_{k,i} * x_i + \theta_1 [WPNC * Yp * x] + \theta_2 [NPWC * Yc * x] +$$
$$\theta_3 [WPWC * Yp * Yc * x] + \varepsilon_i.$$

The usage dummy variable Yp takes on the value of 1 if Actor 1 implements the punishment threat when the option is possible (i.e. WPNC and NPWC), 0 otherwise. Similarly, the usage dummy variable Yc takes on the value of 1 if Actor 1 communicates when the option is possible (i.e., NPWC and WPWC), 0 otherwise. Thus, the coefficient $\theta_1 (\theta_2)$ measures the marginal effect of trust on trustworthiness when the punishment (communication) option is used separately in the WPNC (NPWC) treatment. Likewise, the coefficient θ_3 is a measure of the marginal of trust on trustworthiness when the two mechanisms are implemented simultaneously $(Yp * Yc = 1)$ in the WPWC treatment.

The results of these regressions are summarized in Table 5. Each column summarizes a separate regression, where the back transfer of Actor 2 is served as the dependent variable. The entries in each row are the estimated coefficients, with their standard errors below them in parentheses. The asterisks indicate whether the t-statistics, testing against the hypothesis that the estimated coefficient is zero, is significant at the 10% level (one asterisk), 5% level (two asterisks), or 1% level (three asterisks).

In Model (A), each regression coefficient represents the average back transfer of Actor 2 in each treatment (11.96 in NPNC, 7.11 in WPNC, 9.08 in NPWC, and 7.90 in WPWC). Consistent with the results of our descriptive statistics, the positive and significant coefficients indicate the presence of trustworthiness exhibited by Actor 2 in all four treatments. By dropping the NPNC treatment and re-estimating the regression with a constant term in Model (B), we also learn that the coefficient on the NPNC treatment (11.96) is significantly greater than the coefficients on the WPNC and NPWC treatments (7.11 and 7.90) at $p < 0.05$ and $p < 0.01$ respectively. Following the same approach, the coefficient on the NPNC treatment (9.08) is not significantly different from the coefficient on the NPWC treatment at any conventional level.

This result suggests Actor 2 subjects exhibit more trustworthiness when they are in the condition where Actor 1 has neither the ability to punish nor the ability to communicate, but they exhibit less trustworthiness when the punishment option is available. Our result is also consistent with the findings of previous studies (e.g., Fehr and List, 2004; Houser et al., 2005).

Table 5: Regression Estimates for Actor 2's Back Transfer

Dependent variable: Actor 2's actual back transfer (in shanks)								
	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
Treatment effects								
NPNC	11.96***		0.48					
	(1.40)		(2.57)					
WPNC	7.11***	-4.85***	3.09					
	(1.12)	(1.79)	(1.96)					
NPWC	9.08***	-2.88	-1.13					
	(1.34)	(1.94)	(2.31)					
WPWC	7.90***	-4.06***	0.06					
	(1.10)	(1.78)	(2.17)					
Treatment interaction effects								
x					1.70***			
					(0.14)			
NPNC*x			1.64***	1.70***		1.70***	1.70***	1.70***
			(0.33)	(0.14)		(0.14)	(0.14)	(0.14)
WPNC*x			0.65**	1.05***	-0.65***	1.05***	1.05***	1.05***
			(0.28)	(0.13)	(0.19)	(0.19)	(0.13)	(0.13)
NPWC*x			1.51***	1.38***	-0.32*	0.88***	0.88***	0.88***
			(0.30)	(0.14)	(0.18)	(0.27)	(0.27)	(0.26)
WPWC*x			1.17**	1.18***	-0.52***	1.41***	1.18***	1.41***
			(0.30)	(0.12)	(0.19)	(0.19)	(0.12)	(0.19)
Usage effects								
WPNC *Yp*x						-0.01		
						(0.26)		
NPWC*Yc*x						0.69**	0.69**	0.69**
						(0.31)	(0.31)	(0.31)
WPWC*Yp*Yc*x						-0.38		-0.38
						(0.24)		(0.24)
Constant		11.96***						
		(1.40)						
R²	0.63	0.06	0.77	0.77	0.40	0.79	0.77	0.78
F-Statistic	52.73	2.70	49.64	99.60	20.39	60.01	83.17	70.60
Observations	126	126	126	126	126	126	126	126

Notes: Absolute values of standard errors are in parentheses.

    ***Significant at 0.01 level

    **Significant at 0.05 level

    *Significant at 0.10 level

Model (C) includes the interaction terms between the treatment dummy variables and the transfers of Actor 1 (i.e., NPNC*x, WPNC*x, NPWC*x, and WPWC*x). The estimated coefficients on the treatment dummy variables all become indistinguishable from zero when the interaction terms are incorporated in the regression. This implies the availability of punishment and communication does not directly impact Actor 2's trustworthiness in an absolute sense, but rather it tends to affect the marginal effect of trust on trustworthiness.

We suspect that the treatment dummy might be an irrelevant variable in the model, which can potentially cause an increase in the variance of the estimator. To test this hypothesis and to avoid the problem of adding irrelevant variable, we then estimate Model (D) by excluding the treatment dummy variables that are not significantly different from zero in Model (C). As shown in Table 5, the coefficients on the interaction terms in Model (D) are still statistically significant different from zero, and the R-squared value (0.77) is not sensitive to the exclusion of the treatment dummy variables. Therefore, we confirm that the treatment dummy is irrelevant and can be removed from the model.

Model (D) also suggests Actor 2 in the baseline NPNC treatment reciprocates the trusting behavior of Actor 1 by returning 1.70 shanks for each shank received. This marginal effect of trust on trustworthiness tends to decline significantly when adding the ability to punish or the ability communicate (1.05 in WPNC, 1.38 in WPNC, and 1.18 in WPWC). To determine the level of statistical difference, we re-estimate the regression by replacing the NPNC*x variable with the transfer x variable in Model (E). The results

suggest the coefficient on the NPNC*x variable is significantly greater than the coefficients of other interaction terms.

Model (F) is similar to Model (E) but allows for the interaction variables (i.e., WPNC*Yp*x, NPWC*Yc*x, and WPWC*Yp*Yc*x) to capture the usage effects of punishment, communication, and both mechanisms on trustworthiness. The model indicates only the coefficient on the NPWC*Yc*x variable (0.69) is positive and significant ($p < 0.05$), while the coefficients on other usage interaction variables are negative and not significant at any conventional level (-0.01 for WPNC*Yp*x and -0.38 for WPWC*Yp*Yc*x).

This result provides evidence that a written message from Actor 1 to Actor 2 is not a cheap talk as predicted by the self-interest hypothesis, but it is an effective mechanism to reinforce the ability of trust to elicit trustworthiness. Implementing the punishment threat does not substantially crowds out the ability of trust to elicit trustworthiness. The coefficient on the WPNC*Yp*x variable is negative but it is not significant. Thus, we are unable to replicate the findings of Fehr and List (2004) that not using a punishment threat elicits trustworthiness.

We estimate models (G) and (H) by excluding the usage interaction variables that are not statistically significant in Model (F). As shown in Table 5, the coefficient on the NPWC*Yc*x variable (0.69) is insensitive to the exclusion and remains statistically significant at $p < 0.05$ in both regression models. This confirms the effectiveness of using the communication option. It enhances the marginal effect of trust on trustworthiness.

Model (G) indicates the marginal effect of trust on trustworthiness is highest when neither the ability to punish nor the ability to communicate is available ($\beta_1 = 1.70$ for NPNC). Or, conversely, the availability of punishment and communication options reduces the marginal effects of trust on trustworthiness ($\beta_2 = 1.05$ for WPNC; $\beta_3 = 0.88$ for NPWC; and $\beta_4 = 1.18$). Model (G) indicates communication should be used when it is available. When Actor 1 in the NPWC treatment sends a message to Actor 2, the marginal effect of trust on trustworthiness increase from 0.88 to 1.57 ($0.88 + 0.69 = 1.57$).

In summary, our regression results from models (A) – (H) indicates the presence of both the ability to punish and the ability to communicate can reduce the marginal effect of trust on trustworthiness. However, this marginal loss can be mitigate if Actor 1 uses the communication option. Figure 1 presents the predicted back transfer y, or the predicted level of trustworthiness. The positive impact of trust on trustworthiness is shown. A larger transfer x yields a larger back transfer y in each case. Trust is most productive in the standard NPNC trust game where no punishment and no communication can occur. The larger marginal effect of using communication when it is available is also evident.[5]

---

[5] Also note that subjects in the experiment are allowed to play the trust game twice. That is, after the first game, Actor 1 and Actor 2 are then randomly paired and reassigned to play the second game in one the four experimental treatments that they never experience before. Thus, a great deal of care has been taken to ensure that there is no significant learning effect between the first and the second games. The results obtained from a variety of econometric tests strongly suggest the absent of learning effect and the consistency of behavior (See. Appendix C).

Figure 1: Relationship between Transfer and Actual Back Transfer
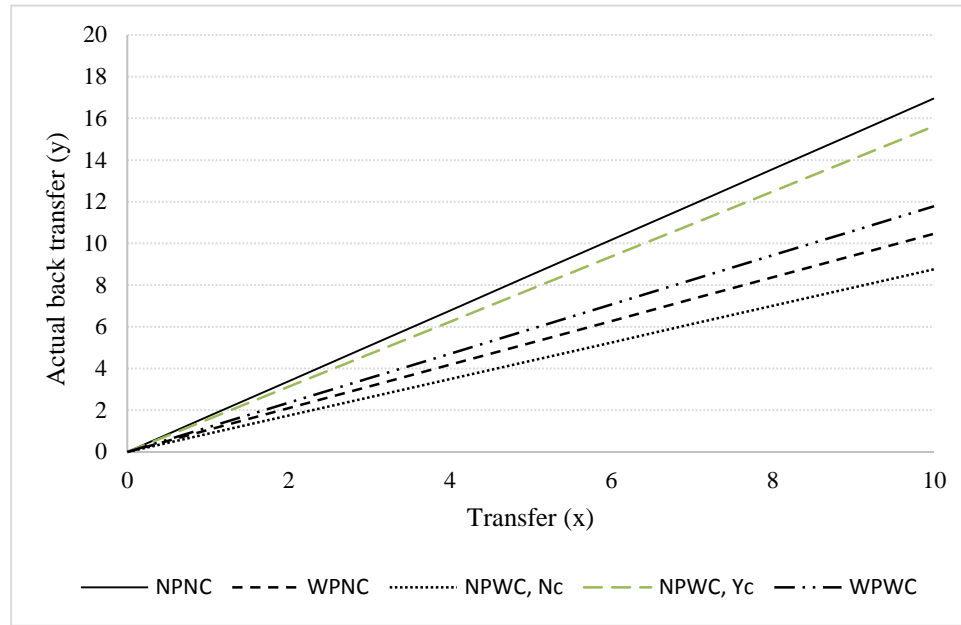
5. Conclusion

In this study, we conducted a trust game experiment with four treatments to obtain improved understanding about how individuals respond to the availability of punishment and communication. The NPNC treatment is the standard trust game, providing a baseline measure of trust and trustworthiness. The WPNC treatment allows Actor 1 to penalize trustee for not providing the desired back transfer, whereas the NPWC treatment allows Actor 1 to communicate with trustee in the form of a written message. The WPWC treatment combines the WPNC treatment with the NPWC treatment, providing Actor 1 with both the ability to punish and the ability communicate.

Consistent with previous studies, we find the presence of trust and trustworthiness in all four treatments. Subjects are not motivated purely by self-interest motive. We also find that communication in the form of a one-way written message by Actor 1 is not

cheap talk, rather it has a significant positive effect on eliciting Actor 2's trustworthiness. However, we did not find a significant positive effect of refraining from implementing the punishment, as suggested by previous studies.

A central findings from the treatments is the availability of punishment and communication significantly lessens marginal effect of trust on trustworthiness. Actor 1 can mitigate this loss associated with communication by choosing to use the option to communicate. Our results also indicate that, if both options are used, the choice to punish tends to cancel out the positive effect which communication tends to have when it is used separately. The vast majority of subjects in the WPWC experiment tend to apply both of mechanisms, not anticipating the reduction in fruitfulness of their trust. Future research could seek to understand this phenomenon.

There are a number of possible extensions to our study. First, we find one-way communication from the trustor is not a cheap talk, but rather it can promote trsuworthy behavior. However, not all forms of communication are effective. Future work could provide insight about what types of messages help translate trust into trustworthiness and why.

Second, a large literature argues generous intention is a determinant of positive reciprocity (e.g., Falk et al., 2008; McCabe et al., 2003; Xiao and Bicchieri, 2010). This approach suggests people are kind to the other if they feel that the other intends to be kind. In our experimental setting, Actor 1 can express his intention through the desired back transfer. Of particular interest is the possibility that Actor 2 appreciates Actor 1 who intends to split earnings equally, so this intention is rewarded with positive reciprocity. This is an interesting task for future research.

Finally, given the body of experimental evidence which systematically refutes the self-interest hypothesis, future research is needed to advance theoretical alternatives to self-interest. The multiplicity of alternative (i.e., altruism, fairness, inequality aversion, positive reciprocity, and intention appreciation) suggest the need for a meta theory.

6. References

Allert, J. R., & Chatterjee, S. R. (1997). Corporate Communication and Trust in Leadership. *Corporate Communications: An International Journal*, 2(1), 14-21.

Andreoni, J., & Miller, J. (2002). Giving According to GARP: An experimental test of the rationality of altruism. *Econometrica*, 70(2), 737-753.

Bacharach, M., Guerra, G., & Zizzo, D. J. (2007). The Self-Fulfilling Property of Trust: An experimental study. *Theory and Decision*, 63(4), 349-388.

Barr, A. (2003). Trust and Expected Trustworthiness: Experimental evidence from Zimbabwean villages. *The Economic Journal*, 113 (489), 614-630.

Bellemare, C., & Kroger, S. (2007).On Representative Social Capital. *European Economic Review*, 51(1), 183-202.

Ben-Ner, A., Putterman, L., & Ren, T. (2011). Lavish Returns on Cheap Talk: Two-way communication in trust games. *Journal of Socio-Economics*, 40(1), 1-13.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior,* 10(1), 122-142.

Bicchieri, C., Lev-On, A., & Chavez, A. (2010). The Medium or the Message? Communication richness and relevance in trust games. *Synthese*, 176(1), 125-147.

Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., & Pancotto, F. (2013). Cooperation Hidden Frontiers: The behavioral foundations of the Italian North-South divide. *Working Papers wp882*, Dipartimento Scienze Economiche, Universita' di Bologna.

Bohnet, I., & Baytelman, Y. (2007). Institutions and Trust: Implications for preferences, beliefs and behaviour. *Rationality and Society*, 19(1), 99-135.

Bolton, G. E. (1991). A Comparative Model of Bargaining: Theory and evidence. *American Economic Review*, 81(5), 1096-1136.

Bracht, J., & Feltovich, N. (2009). Whatever You Say, Your Reputation Precedes You: Observation and cheap talk in the trust game. *Journal of Public Economics*, 93(9-10), 1036-1044.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Russell Sage Foundation, New York, Princeton University Press.

Carpenter, J., Connolly, C., & Myers, C. (2008). Altruistic Behavior in a Representative Dictator Experiment. *Experimental Economics*, 11(3), 282-298.

Cason, T., Sheremeta, R., & Zhang, J. (2012). Communication and Efficiency in Competitive Coordination Games. *Games and Economic Behavior*, 76(1), 26-43.

Charness, G., Cobo-Reyes, R., & Jiménez, N. (2008). An Investment Game with Third-Party Intervention. *Journal of Economic Behavior and Organization*, 68(1), 18-28.

Charness, G., & Dufwenberg, M. (2006). Promises and Partnership. *Econometrica, Econometric Society,* 74(6), 1579-1601.

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*, 117(3), 817-869.

Cochard, F., Nguyen Van, P., & Willinger, M. (2004). Trusting Behavior in a Repeated Investment Game. *Journal of Economic Behavior & Organization*, 55(1), 31-44.

Cohen, T. R., Wildschut, T., & Insko, C. A. (2010). How Communication Increases Interpersonal Cooperation in Mixed-Motive Situations. *Journal of Experimental Social Psychology,* 46(1), 39-50.

Cooper, D., & Kagel, J. (2009). Other Regarding Preferences: A selective survey of experimental results. Forthcoming in J. H. Kagel, and A. Roth (Eds.), *The Handbook of Experimental Economics (Vol. 2).* Princeton University Press.

Coricelli, G., Morales, L. G., & Mahlstedt, A. (2006). The Investment Game with Asymmetric Information. *Metroeconomica*, 57(1), 13-30.

Cox, J. (2004). How to Identify Trust and Reciprocity. *Games and Economic Behavior*, 46(2), 260-281.

Cox, J. C., Sadiraj, K., & Sadiraj, V. (2001). *Trust, Fear, Reciprocity and Altruism*. Mimeo. University of Arizona.

Dirks, K. T., & Ferrin, D. L. (2002). Trust in Leadership: Meta-analytic findings and implications for research and practice. *Journal of Applied Psychology*, 87(4), 611-628.

Duffy, J., & Feltovich, N. (2002). Do Actions Speak Louder Than Words? An experimental comparison of observation and cheap talk. *Games and Economic Behavior*, 39(1), 1-27.

Eckel, C., & Grossman, P. (2008). Differences in the Economic Decisions of Men and Women: Experimental evidence. In C. Plott and V. Smith, V. (Eds.), *Handbook of Experimental Economics Results, pp. 509-519.* New York: Elsevier.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing Theories of Fairness: Intentions matter. *Games and Economic Behavior*, 62(1), 287-303.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a Contract Enforcement Device: Experimental evidence. *Econometrica*, 65(4), 833-860.

Fehr, E., & List, J. (2004). The Hidden Costs and Returns of Incentives-Trust and Trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), 743-771.

Fehr, E., & Rockenbach, B. (2003). Detrimental Effects of Sanctions on Human Altruism. *Microeconomics 0305007*, EconWPA.

Fehr, E., & Schmidt, K. (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics,* 114(3), 817-851.

Frank, R. H. (1988). Passion within Reason: The strategic role of the emotions. New York, NY: W.W. Norton & Company, Inc.

Fukuyama, F. (1995). *Trust*. New York: Free Press.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring Trust. *Quarterly Journal of Economics.* 115(3), 811-846.

Houser, D., Xiao, E., McCabe, K., & Smith, V. (2008). When Punishment Fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior*, 62(2), 509-532.

Isaac, M., & Walker, J. (1998). Nash as an Organizing Principle in the Voluntary Provision of Public Goods: Experimental evidence. *Experimental Economics*, 1(3), 191-206.

Johnson, N. & Mislin. A. (2011). Trust Games: A meta-analysis. *Journal of Economic Psychology,* 32(5), 865-889.

Kimbrough, E. O., Smith, V. L., & Wilson, B. J. (2008). Historical Property Rights, Sociality, and the Emergence of Impersonal Exchange in Long-Distance Trade. *American Economic Review*, 98(3), 1009-1039.

Knack, S., & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A cross-country investigation. *Quarterly Journal of Economics*, 112(4), 1251-1288.

Kottila, M., & Ronni, P. (2008). Collaboration and Trust in Two Organic Food Chains. *British Food Journal*, 110 (4/5), 376-394.

Massey, G., & Kyriazis, E. (2007). Interpersonal Trust between Marketing and R&D during New Product Development Projects. *European Journal of Marketing*, 41 (9/10), 1146-1172.

McCabe, K. A., Rigdon, M. L., & Smith, V. L. (2003). Positive Reciprocity and Intentions in Trust Games. *Journal of Economic Behavior and Organization*, 52(2), 267-275.

Naef, M., & Schupp, J. (2009). Measuring Trust: Experiments and surveys in contrast and combination. *IZA Discussion Papers 4087*, Institute for the Study of Labor (IZA).

Ortmann, A., Fitzgerald, J., & Boeing, C. (2000). Trust, Reciprocity, and Social History: A re-examination. *Experimental Economics*, 3(1), 81-100.

Putnam, R. (1993). *Making Democracy Work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5), 1281-1302.

Rigdon, M. (2009). Trust and Reciprocity in Incentive Contracting. *Journal of Economic Behavior and Organization*, 70(1-2), 93-105.

Sapienza, P., Toldra Simats, A., & Zingales, L. (2007). Understanding Trust. CEPR *Discussion Papers 6462*, C.E.P.R. Discussion Papers.

Schotter, A., & Sopher, B. (2006). Trust and Trustworthiness in Games: An experimental study of intergenerational advice. *Experimental Economics*, 9(2), 123-145.

Sheremeta, R., & Zhang, J. (2013). Three-Player Trust Game with Insider Communication. *MPRA Paper 43533*, University Library of Munich, Germany.

Webster, J., & Wong, W. (2008). Comparing Traditional and Virtual Group Forms: Identity, communication and trust in naturally occurring project teams. *International Journal of Human Resource Management*, 19(1), 41-62.

Willinger, M., Keser, C., Lohmann, C., & Usunier, J.-C. (2003). A Comparison of Trust and Reciprocity between France and Germany: Experimental investigation based on the investment game. *Journal of Economic Psychology*, 24(4), 447-466.\

Wilson, R. K., & Eckel, C. C. (2010). Trust and Social Exchange. In J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science*. Boston: Cambridge University Press.

Xiao, E., & Bicchieri, C. (2010). When Equality Trumps Reciprocity: Evidence from a laboratory experiment. *Journal of Economic Psychology,* 31(3), 456-470.

Appendix

Appendix A: Instructions for the experiment.

You are actor 1

Description of Your Decision Problem

You are a participant in the following decision-making problem. You have been randomly matched with another participant in this problem who is in another room. You will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity. You are in the role of actor 1 and the matched participant is in the role of actor 2. You as well as actor 2 participate only once in this decision problem. You make your decisions with the help of the decision sheet that has been handed out together with this description. Here are the rules that you and actor 2 have to obey when you make your decisions:

Endowment

At the beginning both actors receive an initial endowment of 10 shanks (experimental currency units)

Your decision

You have to make a decision that consists of two components:

1) A *transfer* between 0 and 10 shanks <u>to</u> actor 2.

> You can transfer any amount between 0 and 10 shanks **to actor 2**. You make this decision by indicating a number between 0 and 10 in the appropriate box on your decision sheet. We will then triple this transferred amount, i.e., actor 2 receives three times the amount of shanks you transferred.

2) A *desired back-transfer* <u>from</u> actor 2.

> After you have made your transfer to actor 2 you indicate a desired back-transfer on your decision sheet. The desired back-transfer is the amount you would like to receive back from actor 2. The desired back-transfer can be any number between 0 and three times the amount you have transferred.

The decision of actor 2

Once you have fixed both components of your decision sheet, we collect your decision sheet and give it to actor 2. In this way we inform actor 2 about your decisions. The actor 2 can transfer any amount of the total number of shanks he received **back to you**.

Payoffs

You as actor 1 received: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

Actor 2 receives: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1.

<u>Exchange rate</u>: For every shank you earn you will be paid $2 (2 U.S. dollars).

Description of a New Decision Problem

You now participate in a new decision-making problem. As before, you have been randomly matched with another participant in another room. You are again in the role of **actor 1**. The other participant is in the role of actor 2. Notice that in this new decision problem you are matched with a **new** person, i.e., actor 2 is now a different person compared too the previous problem. Once again, you will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity.

The new decision problem is—with one exception—identical to the previous problem. The exception concerns the conditional payoff cut. ***In the new problem you can impose a condition payoff cut of 4 shanks on actor 2.**** In every other respect the problem is the same. Thus both actors again receive an initial endowment of 10 shanks.

Your decision

Again you have to indicate on your decision sheet what amount you want to transfer to actor 2 and what your desired back-transfer is. Actor 2 receives three times the amount of shanks you transferred.

In addition to the transfer and desired back-transfer you also have to indicate on your decision sheet if you want to impose a conditional payoff cut of 4 shanks on actor 2.

- A conditional payoff cut of 4 shanks for actor 2 has the following consequences: The payoff of actor 2 will be reduced by 4 shanks if his actual back-transfer is less than your desired back-transfer. The conditional payoff cut is <u>not</u> due, i.e., it does not reduce the income of actor 2, if actor 2 transfers exactly your desired amount or more to you
- If you do not impose a conditional payoff cut—the income of actor 2 will not be reduced, irrespective of how large the back-transfer of actor 2 is.

The decision of actor 2

Once you have fixed all three components of your decision sheet, we collect your decision sheet and give it to actor 2. In this way we inform actor 2 about your decisions. The actor 2 can transfer any amount of the total number of shanks he received **back to you**. In case that you have chose a conditional payoff cut of 4 shanks, and if actor 2 transfers back less than what you desired, the conditional payoff cut is due.

Payoffs

You as actor 1 received: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

Actor 2 receives: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1 - 4 shanks (in case that a conditional payoff cut has been imposed **and** is due).

<u>Exchange rate</u>: For every shank you earn you will be paid $2 (2 U.S. dollars).

Description of Your Decision Problem

You are a participant in the following decision-making problem. You have been randomly matched with another participant in this problem who is in another room. You will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity. You are in the role of **actor 2** and the matched participant is in the role of actor 2. You as well as actor 1 participate only once in this decision problem. You make your decisions with the help of the decisions sheet that has been handed out together with this description. Here are the rules that you and actor 1 have to obey when you make your decisions:

Endowment

At the beginning both actors receive an initial endowment of 10 shanks (experimental currency units).

The decision of actor 1

First actor 1 has to make a decision that consists of the following two components:

1) A *transfer* between 0 and 10 shanks <u>to</u> actor 2.

> Actor 1 can transfer any amount between 0 and 10 shanks **to you**. Actor 1 makes this decision by indicating a number between 0 and 10 in the appropriate box on the decision sheet. We will then triple this transferred amount, i.e., you will receive three times the amount of shanks you transferred.

2) A *desired back-transfer* <u>from</u> actor 2.

> After actor 1 has made a transfer to you he indicated a desired back-transfer on your decision sheet. The desired back-transfer is the amount he would like to receive back from you. The desired back-transfer can be any number between 0 and three times the amount that actor 1 has transferred to you.

Your decision

Once actor 1 has fixed both components of the decision, we collect the decision sheet and give it to you. In this way we inform you about actor 1's decisions. Then you can transfer any amount of the total number of shanks you received **back to actor 1**.

Payoffs

Actor 1 receives: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

You as actor 2 receive: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1.

Exchange rate: For every shank you earn you will be paid $2 (2 U.S. dollars).

Description of a New Decision Problem

You will now participate in a new decision-making problem. As before, you have been randomly matched with another participant in another room. You are again in the role of **actor 2**. The other participant is in the role of actor 1. Notice that in this new decision problem you are matched with a **new person**, i.e., actor 1 is now a different person compared to the previous problem. Once again, you will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity.

The new decision problem is—with one exception—identical to the previous problem. The exception concerns the conditional payoff cut. ***In the new problem actor 1 can impose a condition payoff cut of 4 shanks on you.*** In every other respect the problem is the same. Thus both actors again receive an initial endowment of 10 shanks.

The decision of actor 1

Again actor 1 has to indicate on the decision sheet what amount he wants to transfer to you and what his desired back-transfer is. You receive three times the amount of shanks actor 1 transferred to you.

In addition to the transfer and desired back-transfer actor 1 also has to indicate on the decision sheet if you want to impose a conditional payoff cut of 4 shanks on you.

- A conditional payoff cut of 4 shanks has the following consequences for you: Your payoff will be reduced by 4 shanks if your actual back-transfer is less than the desired back-transfer of actor 1. The conditional payoff cut is <u>not</u> due, i.e., it does not reduce your income, if you transfer exactly the desired amount or more to actor 1.
- If actor 1 does not impose a conditional payoff cut—your income will not be reduced, irrespective of how large your back-transfer to actor 1 is.

Your decision

Once actor 1 has fixed all three components of the decision, we collect the decision sheet and give it to you. In this way we inform you about actor 1's decisions. Then you can transfer any amount of the total number of shanks received **back to actor 1**. In case that actor 1 imposed a conditional payoff cut of 4 shanks, and if you transfer back less than actor 1's desired amount, the conditional payoff cut is due.

Payoffs

Actor 1 receives: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

You as actor 2 receive: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1 - 4 shanks (in case that a conditional payoff cut has been imposed **and** is due).

Exchange rate: For every shank you earn you will be paid $2 (2 U.S. dollars).

Appendix B: Demographic questions

Individual demographic questions

Race	White, Non-White
Sex	Male, Female
Did your father obtain a college degree?	Yes, No
Did your mother obtain a college degree?	Yes, No
Do you have a college degree?	Yes, No
Did you complete high school?	Yes, No
Did your parents remain married while you grew up?	Yes, No
Were you a victim of some significant kind of abuse as a child?	Yes, No
Are you married?	Yes, No
Are you an only child?	Yes, No
Size of your income relative to others at the university	Average, Below average, Above average
Size of city where you grew (best representation)	Small town/rural, Small City, Big City

Nationality
            _____
Religion
            _____

Your age	_____
Number of close friends	_____
Number of hours worked per week for pay	_____
Number of volunteer organizations you belong to	_____
Number of hours spent volunteering per week	_____

"Attitude" Questions (or opinions that might have a bearing upon trust or trustworthiness)

Questions from the General Social Survey

> Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?
>
>   Most people can be trusted   Can't be too careful

> Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?
>
>   Would take advantage of you   Would try to be fair

> Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?
>
>   Try to be helpful   Looking out for themselves

Other Questions

To what extent do you agree with the following statements:

"Personal income should not be determined by work."

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"You cannot count on strangers anymore."

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"I trust others."

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"You cannot trust strangers anymore"

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"I am trustworthy."

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"When dealing with strangers, one is better off using caution before trusting them"

Strongly agree   Somewhat agree   Undecided   Somewhat disagree   Strongly disagree

"Behavior" questions (or questions about past actions involving trust)

How often do you leave a door of yours unlocked?

Very Often   Often   Sometimes   Rarely   Never

How often do you lend money to others?

Very Often   Often   Sometimes   Rarely   Never

How often do you lend personal possessions other than money to others?

Very Often   Often   Sometimes   Rarely   Never

How often have you benefited from the generosity of a person you did not know?

Very Often   Often   Sometimes   Rarely   Never

Honesty Index Questions

How much have you lied to your parents?

Very Often   Often   Sometimes   Rarely   Never

How much have you lied to your close friends?

Very Often   Often   Sometimes   Rarely   Never

How much have you lied to acquaintances?

     Very Often   Often   Sometimes   Rarely   Never

How much have you lied to strangers?

     Very Often   Often   Sometimes   Rarely   Never

Appendix C: Testing for learning effects

To examine the presence of learning effect during the experiment, we follow two different approaches including a Wald chi-square test of equality and game-specific dummy variable.

A Wald Chi-Square Test of Equality

First, we estimate a separate regression for Game 1 and Game 2 using Model (G), which is the best fit regression model:

$$y = \sum_{k=1}^{4} \gamma_k T_k x + \theta_2 T_3 Y_c x + \varepsilon = \gamma_1 T_1 x + \gamma_2 T_2 x + \gamma_3 T_3 x + \gamma_4 T_4 x + \theta_2 T_3 Y_c x + \varepsilon$$

Then, we examine whether the estimated coefficients differ across the two games using the Wald chi-square test.

Dependent variable: Actor 2's actual back transfer (in shanks)				
	Game 1	Game 2	Chi-Square χ^2	Probability
NPNC * x	2.01***	1.50***	6.21	0.23
	(0.24)	(0.21)		
WPNC * x	1.05***	1.55***		
	(0.16)	(0.26)		
NPWC * x	0.50	0.94***		
	(0.71)	(0.34)		
WPWC *x	1.36***	1.06***		
	(0.20)	(0.18)		
NPWC*Yc*x	1.24*	0.19		
	(0.74)	(0.49)		
R^2	0.82	0.75		
Observations	63	63		

Notes: Null hypothesis: No behavioral difference between Game 1 and Game 2; $\gamma_k^{Game1} = \gamma_k^{Game1}$, and $\theta^{Game1} = \theta^{Game2}$. Absolute values of standard errors are in parentheses.
      ***Significant at 0.01 level
       **Significant at 0.05 level
        *Significant at 0.10 level

The result from the Wald test chi-square test ($\chi^2 = 6.21$) indicates that there is no significant difference in the estimated coefficients between Game 1 and Game 2. Thus, we cannot reject the

null hypothesis, and conclude that allowing subjects to play the game twice does not create any learning effect in the experiment.

Game-Specific Dummy Variable

This approach allows a dummy variable to capture behavioral difference between Game 1 and Game 2. We create the dummy variable G_2 to categorize the second game: $G_2 = 1$ if subject plays the game second time, 0 otherwise. Thus, we specify the following regression model.

$$y = \beta G_2 + \gamma_1 T_1 x + \mu_1 T_1 x\, G_2 + \gamma_2 T_2 x + \mu_2 T_2 x G_2 + \gamma_3 T_3 x + \gamma_3 T_3 x G_2 + \gamma_4 T_4 x + \gamma_4 T_4 x G_2$$
$$+ \theta T_3 Y_c x + \varphi T_3 Y_c x G_2 + \varepsilon$$

The coefficient β indicates the different level of Actor 2's back transfer between Game 1 and Game 2. The coefficient μ_i indicates the difference in the marginal effect of trust in each treatment between the two games, and the coefficient φ indicates how the use of communication in Game 1 differs from Game 2.

Dependent variable: Actor 2's actual back transfer (in shanks)		
	Coef.	Std. Err.
NPNC * x	2.01***	0.23
NPNC * x * G_2	-0.53	0.33
WPNC * x	1.04***	0.15
WPNC * x * G_2	-0.02	0.36
NPWC * x	0.50	0.67
NPWC * x * G_2	0.42	0.74
WPWC * x	1.35***	0.18
WPWC * x * G_2	-0.31	0.30
NPWC* Yc * x	1.24**	0.69
NPWC* Yc * x * G_2	-1.05	0.81
G_2	0.21	1.45
R^2	0.78	
Observations	126	

Notes: Absolute values of standard errors are in parentheses.
***Significant at 0.01 level
**Significant at 0.05 level
*Significant at 0.10 level

The results indicate all of the estimated coefficients on the dummy variables G_2 are not statistically significant at any conventional level. Thus, subjects do not behave differently between Game 1 and Game 2, and there is no leaning effect in the experiment.

Chapter 3

Eliciting Positive Reciprocity: The Power of Proposing Equality

Abstract

This paper presents a theoretical model of play in the trust game that specifically features the potential which proposing equality has for promoting positive reciprocal behavior. The model proposes the utility function where the second mover appreciates the first mover that intends to split all earnings equally, and rewards this intention with positive reciprocity. The second mover optimizes given the preference, and the first mover optimizes while recognizing the optimizing behavior and preference of the second mover. The optimizing behavior is then tested econometrically against the trust experimental data in which the first mover has an opportunity to numerically indicate the desired amount that he would like to receive back from the second mover. The findings suggest that the second mover maximizes the utility when the first mover proposes equality (two thirds of the tripled investment) and aims to reciprocate this equality intention by sending back the desired amount. On the other hand, when the first mover does not propose equality, the second mover corrects this inequality intention by sending back less than the desired amount.

1. Introduction

The self-interest hypothesis that assumes people are motivated purely by their self-interest has long been a central assumption in simplifying the complexity of human behavior in most economic models. Sear and Funk (1990, p. 248) describe in some detail how the hypothesis is developed:

> "The self-interest hypothesis normally revolves around three basic psychological assumptions: the idea of materialistic hedonism, or a simple pleasure-pain principle of human motivation; the idea of egotism, that outcomes to the self are weighed more heavily than outcomes to others; and the idea of rationality, that decisions are made on the basis of reasonable calculations limited primarily by the amount of information available."

In practice, the self-interest hypothesis has been challenged systematically by the collected evidence from experimental studies. Many economists have proved that in many game experiments subjects do not behave selfishly, but rather they are concerned about morals, fairness, generosity, and the well-being of others (e.g., Charness and Dufwenberg, 2006; Fehr and Gachter, 2002; Fehr and List 2004; Fehr and Rockenbach, 2003).

Such observations in the game experiments have led economists to argue that the self-interest might be a false premise that prevents the ability of economic models to predict behavior. Subsequently, several studies have proposed a new class of theories, known as theories of social preferences, that extends individual satisfaction beyond own material self-interest (e.g., Bolton and Ockenfels, 2000; Cox, Friedman, and Gjerstad, 2006; Charness and Rabin, 2002; Fehr and Schmidt, 1999; Levine, 1998; Rabin, 1993). These newly developed theories indicate people also care about outcomes and actions of others rather than caring exclusively about their own outcomes. Nevertheless, people do

use economic reasoning but not limited to the self-interest, so the concept of utility maximization can still be applied to analyze the behavior and equilibrium of the game experiments.

In the context of a trust game experiment, a large body of experimental evidence suggest subjects frequently exhibit trust and trustworthiness, indicating strong social preferences in the population (e.g., Berg, 1995; Barr, 2003; Glaeser et al., 2000; Cochard, Nguyen and Willinger, 2004; Camerer, 2003; Johnson and Mislin, 2011; Willinger, Keser, Lohmann and Usunier, 2003). Recently, the main focus of the research on this topic has been to apply the existing theories of social preferences to identify a specific type of preference in laboratory experiments (e.g., fairness, altruism, positive reciprocity). To date, however, there is no theoretical model of the trust game that allows for and consolidates multiple types of social preferences in subject's decision making. Thus, the development of this model will provide additional insight into the motives and driving force behind the observed pattern of trusting and trustworthy behavior.

Another aspect of the trust game that has been absent in the literature is the potential of generous intention in promoting positive reciprocity. Previous studies seek to understand subject's intentions and motives using communication in the form of written message, exchange proposal, and face-to-face (e.g., Issac and Walker, 1998; Duffy and Feltovich, 2002; Charness and Dufwenberg, 2006). To our knowledge, no study has recognized the potential of the numerical message, which can be used as a means for the first mover to communicate with the second mover. Most importantly, this form of communication also helps the second mover to recognize the motives and intentions of the first mover in the experiment. The second mover may care about his opponent's intentions. For example, if he feels that the first mover intends to behave generously, he

may want to reward this good intention with positive reciprocity. On the other hand, if he

feels that the first mover intends to behave greedily, he may want to correct this bad

intention with negative reciprocity, or hostile actions.

This study makes three important contributions to the literature. First, the study

introduces a new model of play in the trust game that comprehensively consolidates all types of

social preferences in the second mover's decision making. Second, the model specifically features

the potential which proposing equality has for promoting positive reciprocal behavior. In

particular, we proposes the utility function where the second mover appreciates the first mover

that intends to split all earnings equally, and rewards this intention with positive reciprocity.

Third, this paper illustrates cooperative equilibria in a one-shot trust game may exist if the self-

interested first mover has the ability to recognize the preferences and optimizing behavior of the

second mover.

The remainder of this study is structured as follows. Section 2 provides a more detailed

discussion of recent social preference theories. Section 3 presents our model of social preferences

in the trust game. Section 4 applies this model to the trust game experiment, in which the first

mover has an opportunity to numerically indicate the desired amount that he would like to receive

back from the second mover. Section 5 presents empirical evidence that the majority of subject

actions in the trust game experiment considered in the previous section are consistent with the

prediction of our model. Section 6 provides theory and empirical evidence that the ability to

recognize the second mover's optimizing behavior and preferences may allow the first mover to

exhibit trust. Section 7 summarizes and concludes the paper.

2. Review of Related Theoretical Studies

During the last two decades a number of behavioral economists have attempted to

develop theories to rationalize subjects' behavior in the experiment. These theories generally

reply on the notion of social preferences, assuming people care not only about their own payoffs but also the payoffs allocated to others. More specifically, the utility of person i is sensitive to the variation of payoffs received by person j, x_j, for any given of his resource allocation, x_i.

"Inequity aversion" is a certain type of social preferences, which places an emphasis on the idea that people care for fairness and they are reluctant to incidental inequalities. Fehr and Schmidt (1999) propose the model, where an individual's utility function can either be increasing or decreasing function in the payoffs allocated to others. In the two-person case, the utility of person i is given by

$$1) \quad U_i(x_i, x_j) = x_i - \alpha_i(x_j - x_i, 0) - \beta_i(x_i - x_j, 0)$$

or,

$$U_i(x_i, x_j) = \begin{cases} x_i - \alpha_i(x_j - x_i) & \text{if } x_i < x_j \\ x_i - \beta_i(x_i - x_j, 0) & \text{if } x_i > x_j \end{cases}$$

where $\alpha_i > \beta_i > 0$ and $\beta_i \leq 1$. The model indicates people make decisions so as to minimize inequity in outcomes ($\alpha_i > \beta_i > 0$), and they dislike disadvantageous inequality more than advantageous inequality ($\alpha_i \geq \beta_i$)

Ottone and Ponzano (2005) argue that the linear utility function is a critical drawback in the Fehr-Schmidt model, which prevents interior solutions in some experimental games. To address the problem, Ottone and Ponzano proposes a further modification to the model by allowing the concavity in the utility function. Their model is given by

$$2) \quad U_i(x_i, x_j) = x_i - \alpha_i \left(\frac{x_j - x_i}{\sigma x_i + 1}, 0 \right) - \beta_i \left(\frac{x_i - x_j}{\gamma x_i + 1}, 0 \right)$$

The model works in the exact same manner as the original Fehr-Schmidt model, except assumes that people are concerned about their incomes as an unequal weight difference among population. This unequal weight is categorized by the parameters $0 < \sigma$ and $\gamma < 1$.

The alternative model of inequity aversion is independently developed by Bolton and Ockenfels (2000). The model has provided substantial success in rationalizing a variety of puzzling evidence in many experiments (i.e., generosity in dictator, trust game, and gift exchange). In the model, the utility function of person i is given by

3) $U_i = U_i(x_i, \sigma_i)$

and

$$\sigma_i = \begin{cases} \frac{x_i}{\sum_{j=1}^{N} x_j} & if \ \sum_{j=1}^{N} x_j \neq 0 \\ \frac{1}{N} & if \ \sum_{j=1}^{N} x_j = 0 \end{cases}$$

where x_i is person i's own payoff and σ_i is person i's share of income. For any given x_i, the utility function of person i is strictly increasing and diminishing in the share of income σ_i.

Unlike Fehr and Schmidt, the Bolton-Ockenfels model assumes that person i is more concerned about the average income of all people than the relative income of the opponent. Note that $dU_i/d\ x_j$ does not depend on person j's relative position toward person i, but on how greater person i does when he compares himself to the average. That is, if person i stands at the position where his income x_i is above the average, he is willing to allocate more resources to person j regardless of whether person j is at a better or poorer position. On the other hand, if x_i is below the average, person i would reduce person j's income irrespective of whether person j has a much lower income.

"Altruism" is another type of social preference, which has been used extensively to describe the pattern of behavior observed during experiments. Behavioral economists formally define altruism as an unconditional form of kindness, or the willingness to sacrifice own resources in order to promote the well-being of others without an expectation of receiving any recognition from the act (Andreoni, 1989; Cox et al., 2001; Andreoni and Miller, 2002; Charness and Rabin, 2002; and Fehr and Schmidt, 2006).

The simplest version of the altruistic model only includes a parameter to express a concern for the overall efficiency, so

4) $U_i(x_i, \dots, x_N) = x_i + \sum_{i \neq j} \gamma_i x_j \quad where \; \gamma_i > 0$

The model indicates, for an altruist, his or her utility function is positive and increasing in the well-being of others. The altruist will assign a certain weight on the well-being of all others ($\gamma_i > 0$), so the partial derivatives with respect to the payoffs received by those other people are strictly positive and increasing $dU_i(x_i, \dots, x_N)/dx_j > 0, i \neq j$. Equivalently, this condition implies person i cares about the overall efficiency in addition to his own self-interest.

Andreoni and Miller (2002) conduct a series of dictator game experiments to test the validity and reliability of the simple altruistic model and find that the model is substantially successful at explaining such variations in the data. However, several studies at the time also attempt to develop alternative models of altruism, and the focus of those studies is to connect the Fehr-Schmidt model of inequity aversion to the preference for altruism. (e.g., Levine 1998; Huck et al., 2001; Charness and Rabin, 2002; Erlei, 2004; Cox et al., 2006; Rotemberg, 2004).[6] In the following, we describe some of these models in more detail.

Charness and Rabin (2002) propose the model of quasi-maximin preferences, where the altruistic preference and the Rawlsian inequality aversion together are combined. In the model, the utility function of person i is a convex combination of his own payoff and the social welfare function:

5) $U_i(x_i, \dots, x_N) = (1 - \gamma)x_i + \gamma[\delta * \min\{x_i, \dots, x_N\} + (1 - \delta) * (x_i + \dots + x_N)$

In the two-person case, the model can be written as

---

[6] A comprehensive review for the theoretical models that combine altruism and inequality aversion together is provided in Fehr and Schmidt (2006).

$$U_i(x_i, x_j) = \begin{cases} x_i + \gamma(1 - \delta)x_j & if \ x_i < x_j \\ (1 - \gamma\delta)x_i + \gamma x_j & if \ x_i \geq x_j \end{cases}$$

where δ parameterizes the weight assigned on the maxi-min criterion. The model indicates the altruist cares not only about the well-being of others but also the relative standing of the others.

Similarly, Kohler (2011) argues the models of altruism and inequity aversion both are sharing a similarity in their properties. In particular, he disputes the Fehr-Schmidt model simply ignores the parameter of efficiency concern γ by setting it equal zero. In his work, the utility function is given by

6) $\quad U_i(x_i, x_j) = x_i + \gamma_i x_j - \alpha_i(x_j - x_i, 0) - \beta_i(x_i - x_j, 0)$

Or

$$U_i(x_i, x_j) = \begin{cases} x_i + \gamma_i x_j - \alpha_i(x_j - x_i) & if \ x_j > x_i \\ x_i + \gamma_i x_j - \beta_i(x_i - x_j) & if \ x_j < x_i \end{cases}$$

When $\gamma_i \neq 0$, player i is altruistic and his optimal behavior can be determined directly by a trade-off among the preferences for inequality aversion and altruism (γ, α, β). The model preserves not only the interpretation of the Fehr-Schmidt model but also the preference for altruism in the decision making.

Engelmann (2012) criticizes the simple modification in the work of Kohler (2011) that it provides an implausible prediction for some experimental games. If the altruistic preference is added directly into the inequality aversion model, then as long as experimental games with a given number of players are considered, the parameters of the model cannot be uniquely identified.

In the opposite of the altruistic preference, several theoretical studies seek to explain subjects' behavior by recognizing the ideas of envy and jealousy (Bolton, 1991; Kirchsteiger, 1994) – the willingness to decrease the material payoff of a reference agent at a personal cost.

These studies establish the utility function, where person i's utility is strictly decreasing in person j's payoff $dU_i(x_i, x_j)/dx_j < 0, i \neq j$.

"Reciprocity" is another type of social preferences that has been widely recognized by behavioral economists to explain the pattern of experimental data. Fehr and Gachter (2000) define reciprocity as "an in-kind response to beneficial or harmful acts": people tend to produce rewarding action in response to another positive action and hostile action in response to another negative action.

The theoretical model that incorporates reciprocity into the utility function is originally developed by Cox et al. (2006). The model suggests the following form of the utility function.

$$7) \quad U_i(x_i, x_j) = \begin{cases} \frac{1}{\alpha}(x_i^\alpha + \varphi x_j^\alpha) \ if \ \alpha \neq 0 \\ (x_i * x_j)^\varphi \ if \ \alpha = 0 \end{cases}$$

where α parameterizes the curvature of the indifference curve in the planner of person i's income (x_i) and person j's income (x_j) space . The marginal rate substitution between x_i and x_j is then given by

$$MRS = \frac{dU_i(x_i, x_j)\big/dx_i}{dU_i(x_i, x_j)\big/dx_j} = \varphi^{-1}\left(\frac{x_j}{x_i}\right)^{1-\alpha}$$

The model describes the preferences for altruism and spite throughout the emotional state of person i, which is represented the parameter $\varphi = \varphi(r)$. This emotional state depends upon the reciprocity motive r which is defined as $r(x) = \bar{x}_i(c_j) - x_i^0$, where $\bar{x}_i(c_j)$ is the maximum payoff person i can generate given person j's choice c_j and x_i^0 is an appropriate reference payoff. The reciprocity motive function demonstrates the idea that person i recognizes the additional payoff as person j's kindness, which needs to be reciprocated positively $[\bar{x}_i(c_j) > x_i^0]$ and the

shortfall from his appropriate reference x_i^0 as a violation that needs to be reciprocated

negatively $[\bar{x}_i(c_j) < x_i^0]$.

Another approach in modeling the preference for positive reciprocity is to allow for the

possibility that people may care about their opponent's intention, formally known as "the theories

of intention based reciprocity". For example, an individual may be kind to his opponent if he

believes that his opponent has a generous intention to him.

Rabin (1993) originally develops the theory of the intention based reciprocity for a

simple two-person case by proposing kindness function within the utility function

8) $U_i(a_i, b_j, c_i) = x_i(a_i, b_j) + \bar{f}_i(c_i, b_j)[1 + f_i(a_i, b_j)]$

where $f_i(a_i, b_j)$ denotes the kindness function measuring how kind person i is to person j, and

$\bar{f}_i(c_i, b_j)$ denotes how person i perceives person j's kindness.[7] The first term in the utility function

is simply the standard utility function. In the second term, if person i perceives person j's

kindness $[\bar{f}_i(c_i, b_j) > 0]$, person i receives some additional utility from being kind to person j. On

the other hand, if person j is perceived to be unkind $[\bar{f}_i(c_i, b_j) < 0]$, person i experiences the loss

in his utility and will compensate this loss by being unkind to person j.

The next section provides a theoretical model of play in the trust game, seeking to extract

understanding of how proposing equality has a potential for promoting positive reciprocal

behavior, and how social preferences can influence subject's beliefs and actions.

---

[7] The kindness function of player i to player j is defined as $f_i(a_i, b_j) = \frac{x_j(a_i, b_j) - x_j^e(b_j)}{x_j^h(b_j) - x_j^l(b_j)}$ where a_i denotes strategy chosen by player i and b_j denote player i's belief about what player j is going to do. The variable $x_j^e(b_j)$ is an equitable payoff that player i can give to player j, which is given by $\frac{x_j^h(b_j) - x_j^l(b_j)}{2}$, where $x_j^h(b_j)$ and $x_j^l(b_j)$ are the highest and the lowest payoffs of player j, respectively. Then, the model captures how player I perceive player j's kindness in the same manner, $\bar{f}_i(c_i, b_j) = \frac{x_i(c_i, b_j) - x_i^e(c_j)}{x_i^h(c_i) - x_i^l(c_i)}$ where c_i denotes player i's belief about what player j believes that player i is going to do.

3. The Model

In Berg et al.'s (1995) study of trust and reciprocity, the standard trust game is played as follows. Two subjects are randomly assigned in the role of Actor 1 and Actor 2, and both are paired anonymously without knowing any identity of their matched counterparts. Subjects receive an endowment of 10 shanks (experimental units). Actor 1 chooses a transfer amount of $x \in \{0, 1, 2, \ldots, 10\}$ and indicates a desired back transfer $\hat{y} \in \{0, 1, 2, \ldots, 3x\}$. Each shanks transferred by Actor 1 will triple by the time Actor 2 receive it, and Actor 2 then chooses the actual back transfer $y \in \{0, 1, 2, \ldots, 3x\}$.

The payoff received by Actor 1 is given by $\pi_1 = 10 - x + y$, and the payoff received Actor 2 is given by $\pi_2 = 10 + 3x - y$. The degree of trust in the experiment is captured by the willingness of Actor 1 to transfer a positive amount of initial endowment $(0 < x \leq 10)$ to Actor 2 in the first stage. Likewise, the degree of trustworthiness is the willingness of Actor 2 to send back some amount of tripled transfer $(0 < y \leq 3x)$ in the second stage.

Knowing each shank transferred generates the surplus of three shanks $3x$, Actor 1 who cares for equality will allocate one-third of the total surplus to Actor 2, so that each player can receive the earning of $10 + x$ at the end of the game. Thus, the desired back transfer is twice the amount sent when Actor 1 proposes equality, or $\hat{y} = 2x$. Similarly, if Actor 2 has the preference for equality, he should send back two-thirds of the total surplus $y = 2x$ to Actor 1. [8]

The understanding of these two important conditions allows us to develop a model by recognizing Actor 2 as a combination of the following pure types.

---

[8] These conditions can be determined directly by equalizing the earnings of the two players. $\pi_1 = \pi_2$ implies $10 - x + y = 10 - y + 3x$, or $y = 2x$. Thus, Actor 2 who has preference for equality should set up $y = 2x$ and Actor 1 who proposes equality should also set $\hat{y} = 2x$.

Self -Interested Type: This type of Actor 2 is predicted to behave in the way that is most personally beneficial. Thus, the utility function is given directly by his own payoff received at the end of the game.

9) $u_2^{T1} = 10 - y + 3x$

Since the first derivative with respect to the back transfer is strictly negative $du_2^{T1}/dy = -1$, Actor 2 maximizes his utility by always setting $y = 0$, or exhibiting no trustworthiness.

Inequality Aversion Type: This type of Actor 2 maximizes his utility by always imposing material equality, or setting $y = 2x$. To allow for the global maximum at the equality condition, the utility function for this type of Actor 2 may take the following form

10) $u_2^{T2} = -a_1(y - 2x)^2$

where $a_1 > 0$ parameterizes the degree to which the preference for inequality aversion motivates Actor 2 to exhibit trustworthiness. The rational choice theory indicates $du_2^{T2}/dy = -2a_1[y - 2x]$. Thus, the inequality averse Actor 2 always maximizes his utility at $y = 2x$, and experiences the loss of in his utility whenever Actor 1 has a lower or higher payoff than himself.

Altruism type: This type of Actor 2 is motivated solely by a strong willingness to sacrifice his payoff for the well-being of Actor 1. In technical terms, Actor 2 is altruistic if the first partial derivative of his utility function with respect to Actor 1's payoff is positive and increasing. Thus, an altruistic model can be written as

11) $u_2^{T3} = a_2(y - 2x)D$

and

$$D = \begin{cases} 1 \ if \ y - 2x \geq 0 \ ; \ \pi_2 < \pi_1 \\ 0 \ if \ y - 2x < 0 \ ; \ \pi_2 > \pi_1 \end{cases}$$

where $a_2 > 0$ measures the altruistic preference in Actor 2's decision making. D is a conditional variable that indicates whether Actor 1 or Actor 2 has a better payoff at the end of the game.

Here is how the model works. Actor 2 will choose the optimal back transfer at the level of which his utility is maximized, $du_2^{T3}/dy = a_2 D$. This condition implies whether the marginal utility of Actor 2 will be zero or non-zero is dependent upon the value of the conditional variable D. Since the parameter a_2 is strictly positive for the altruist, Actor 2 receives the additional utility only when Actor 1 achieves a higher payoff: $du_2^{T3}/dy = a_2 > 0; D = 1; \pi_2 < \pi_1$.

Positive reciprocity type: This type of Actor 2 appreciates Actor 1 who exhibits trust, and rewards the trusting behavior by sending a larger back transfer. Technically, the utility function for this type of Actor 2 is strictly increasing in the transfer of Actor 1.

12) $u_2^{T4} = (b_1 x - b_2 x^2)y$

where $b_1 > 0$ indicates the degree of positive reciprocity, and $b_2 > 0$ indicates the possibility of diminishing returns associated with the positive reciprocity. The model indicates the degree of positive reciprocity exhibited by Actor 2 is positive and increasing in the transfer sent by Actor 1, but will start to decline at a certain level. In the work of Rankin and Taborsky (2009), the experimental evidence suggests the diminishing returns in cooperative behaviors are common and widespread in nature.

Since $du_2^{T4}/dy = b_1 x - b_2 x^2$, Actor 2 maximizes his utility by sending back the positive amount only if the transfer of Actor 1 satisfies the following requirements: i) the transfer amount is positive, $x > 0$; and 2) the ratio between the preference for positive reciprocity and the diminishing returns outweighs the transfer amount, $b_1/b_2 > x$. [9]

---

[9] $du_2^{T4}/dy = b_1 x - b_2 x^2 = x(b_1 - b_2 x) > 0$. Therefore, $x > 0$ and $b_1/b_2 > x$

Intention appreciation type: This type of Actor 2 makes an evaluation on his opponent's intention, and then chooses the appropriate level of the back transfer. The utility function can be written as

13) $u_2^{T5} = -b_3(\hat{y} - 2x)^2 y$

where $b_3 > 0$ parameterizes the degree to which Actor 2 cares about Actor 1's intention toward equality. The quadratic term $(\hat{y} - 2x)^2$ implies Actor 2's utility will decline when Actor 1 does not intend to split all earnings equally, irrespective of whether it is advantageous $(\hat{y} - 2x < 0)$ or disadvantageous inequality $(\hat{y} - 2x > 0)$.

Since $du_2^{T4}/dy = -b_3(\hat{y} - 2x)^2$, Actor 2's utility will reach the maximum level whenever Actor 1 intends to generate equality by requesting $\hat{y} = 2x$, and the optimal behavior is to reward this intention by sending back the desired amount.

Mixed type: A mixed type of Actor 2 represents real world population that is motivated by all types of preferences (i.e., self-interest, altruism, inequality aversion, positive reciprocity, and equality intention). Thus, the utility function for this type of Actor 2 is a combination of all pure types.

14) $U_2^{TM} = 10 - y + 3x - a_1(y - 2x)^2 + a_2(y - 2x)D + (b_1 x - b_2 x^2)y + b_3(\hat{y} - 2x)^2 y$

Since $dU_2^{TM}/dy = -1 - 2a_1(y - 2x) + a_2 D - (b_1 x - b_2 x^2) - b_3(\hat{y} - 2x)^2 = 0$, Actor 2 maximizes his utility by setting [10]

15) $y = \dfrac{-1}{2a_1} + \left[\dfrac{b_1 + 4a_1}{2a_1}\right]x - \left[\dfrac{b_2}{2a_1}\right]x^2 - \left[\dfrac{b_3}{2a_1}\right](\hat{y} - 2x)^2 + \left[\dfrac{a_2}{2a_1}\right]D$

---

[10] Appendix A: Mathematical note for theoretical model

Condition (15) indicates a mixed type Actor 2 faces trade-offs because of the multiple types of social preferences. For example, while inequality aversion motivates Actor 2 to set $y = 2x$, self interest motivates a smaller value for y, while altruism motivates a larger value for y. Also, this model assumes Actor 2 will punish an Actor 1 subject for not proposing equality, with the marginal punishment increasing as more inequality is proposed.

In the experiment described in the next section, Actor 1 has an opportunity to indicate a desired back transfer, along with chosen transfer, and Actor 2 then chooses the back transfer. This data is used to test the predictive capability of the mixed model (14). The values of the preference parameters ($a_1, a_2, b_1, b_2,$ and b_3) can be uniquely identified using the least-squares regression to fit the optimization condition (15) to the data.

4. Data and Descriptive Statistics

4.1 Data and Experiment

Our empirical analysis aims to test the ability of the theoretical model in eliciting Actor 2's preferences for Actor 1's equality proposal. We retrieved the experimental data from the work of Wuthisatian et al. (2015) which examines the effects of punishment and communication on trust and trustworthiness.

In the experiment, 260 subjects (130 pairs) recruited from the University of Nevada, Reno, including undergraduate students, graduate student, and employees were randomly assigned into 4 different treatments. The control treatment is the standard trust game, which neither ability to punish nor ability to communicate is available to the first mover. The second treatment is identical to the standard trust condition except Actor 1 has an opportunity to impose the punishment threat upon Actor 2, but does not have the ability to communicate with Actor 2. The third treatment permits Actor 1 to communicate with Actor 2 by writing the message, but the

ability to punish was not allowed. Finally, the forth treatment allows Actor 1 to employ both capabilities.

The experimental data was collected in a series of 10 sessions, where each session randomly assigns subjects into different treatments. For a given session, subjects were randomly allocated into the roles of Actor 1 and Actor 2, except the last few subjects who were placed into a role in order to ensure equal numbers of Actor 1 and Actor 2 players. Actor 1 players were always in one room, while Actor 2 players were in another room down the hall. To ensure anonymity, care was taken, primarily by rapidly shepherding a subject into one room or another, to ensure players of one type did not ever see the players of the other type.

Actor 1 and Actor 2 each receives an endowment of 10 shanks (experimental currency units). By writing on a decision sheet, Actor 1 choses a transfer amount $x \in \{0, 1, 2, \ldots, 10\}$ shanks to send to Actor 2. Actor 1 is allowed to indicate a desired back-transfer $\hat{y} \in \{0, 1, 2, \ldots, 3x\}$. Actor 1 cares about the equality, in this case, his desired back transfer should be $\hat{y} = 2x$. The decision sheet indicated by Actor 1 is then given to the responder Actor 2, with Actor 2 receiving 3 shanks for each shank transferred by Actor 1. Then, Actor 2 choses the actual back-transfer level $y \in \{0, 1, 2, \ldots, 3x\}$. The payoff received by Actor 1 is given by $\pi_1 = 10 - x + y$, whereas the payoff for Actor 2 is $\pi_2 = 10 + 3x - y$. Similarly, if Actor 2 cares about the equality, his back transfer should be given by $y = 2x$. Each participant was paid two U.S. dollars for each shank earned (See Appendix B).

4.2 Descriptive Statistics

Table 1 presents a summary of subject behavior across four experimental treatments. The means of Actor 1's transfer over all treatments is 65% of the endowment, ranging from 61.3% in treatment 3 to 6.77% in treatment 4. There is no significant difference in the average transfer across the four treatments. The average desired back transfer of Actor 1 ranges between 12 and

15 shanks. Interestingly, the desired back transfer of Actor 1 is roughly twice the amount sent in all four treatments. This implies the average Actor 1 subject is proposing to divide the surplus created by his transfer such that the payoffs of the two players are roughly equal. According to the F-test, there is no significant difference (F [3, 126] = 0.14) in the means of the equality intention across the four treatments.

Table 1: A Summary of Subjects' Behavior

Variable	All subjects	Treatment 1	Treatment 2	Treatment 3	Treatment 4
	Avg.	Avg.	Avg.	Avg.	Avg.
x	6.50	6.72	6.13	6.77	6.55
	(3.37)	(3.68)	(3.33)	(3.65)	(3.11)
y	9.00	11.88	8.10	9.08	8.03
	(7.42)	(7.49)	(7.63)	(8.33)	(6.25)
$\hat{y}$	13.42	14.32	12.49	13.81	13.53
	(7.60)	(8.04)	(7.54)	(8.20)	(7.16)
$y - 2x$	-4.00	-1.56	-4.15	-4.46	-5.08
	(6.86)	(5.66)	(8.14)	(6.51)	(6.26)
$\hat{y} - 2x$	0.42	0.88	0.23	0.27	0.43
	(4.06)	(3.38)	(4.43)	(4.04)	(4.20)
Observations with $\pi_2 < \pi_1$	77	10	25	13	29
Observations	130	25	39	26	40

Notes: Standard deviations (S.D.) are in parentheses.

**Figure 1:** Histogram of Proposing Equality
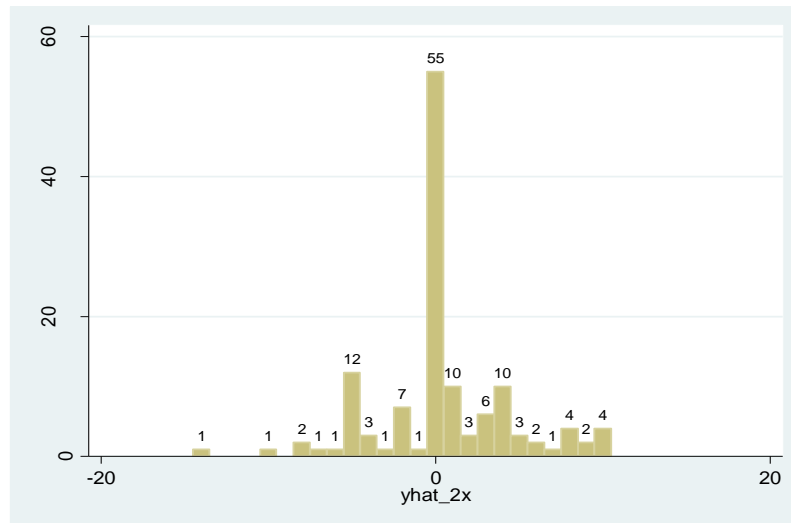
The histogram presented in Figure 1 indicates 42% of Actor 1 participants (55 of 130 subjects) express a concern for fair outcomes by proposing to split the earnings of the two players equally. We also find that 35% of Actor 1 participants (45 out of 130 subjects) intend to earn more money than Actor 2 by requesting back above the equality condition, and only 23 % of Actor 1 participants (30 out of 130 subjects) intend to earn less money than Actor 2 by requesting back below the equality condition.

To further examine how Actor 2 responds to Actor 1's equality proposal, we separate the data of each treatment according to whether Actor 1 intends to split all earnings equally. As shown in Table 2, the average back transfer of Actor 2 is significantly higher ($p < 0.05$) in every experimental treatment when Actor 1 proposes equality (12.64 in Treatment 1, 8.50 in Treatment 2, 11.93 in treatment 3, and 9.54 in treatment 4). This result provides a strong indication that Actor 2 appreciates Actor 1 who intends to split all earnings equally, and reward this intention with a higher degree of positive reciprocity and trustworthiness. Further support for this difference in Actor 2's trustworthiness can be seen in Figure 2.

Table 2: Splitting Out Proposing Equality

Variable	All subject		Treatment 1		Treatment 2		Treatment 3		Treatment 4	
	$\hat{y} = 2x$	$\hat{y} \neq 2x$	$\hat{y} = 2x$	$\hat{y} \neq 2x$	$\hat{y} = 2x$	$\hat{y} \neq 2x$	$\hat{y} = 2x$	$\hat{y} \neq 2x$	$\hat{y} = 2x$	$\hat{y} \neq 2x$
x	6.84	6.25	6.79	6.64	6.79	5.76	6.36	7.25	7.46	6.11
	(3.87)	(2.96)	(4.28)	(2.94)	(3.72)	(3.11)	(4.60)	(2.18)	(2.96)	(3.14)
y	10.67	7.77	12.64	10.91	8.50	7.88	11.93	5.75	9.54	7.30
	(7.57)	(7.10)	(7.38)	(7.88)	(6.53)	(8.30)	(9.17)	(6.00)	(6.97)	(5.88)
$\hat{y}$	13.67	13.24	13.57	15.27	13.57	11.88	12.71	15.08	14.92	12.85
	(7.73)	(7.55)	(8.56)	(7.62)	(7.45)	(7.67)	(9.20)	(7.03)	(5.92)	(7.70)
Observations	55	75	14	11	14	25	14	12	13	27

Notes: Standard deviations (S.D.) are in parentheses.

Figure 2: Average Back Transfer by Proposing Equality
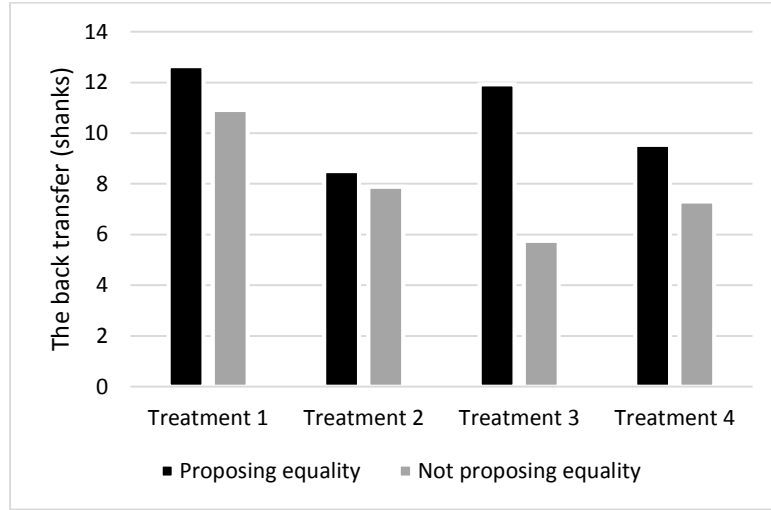
5. Empirical Results

To measure the ability of our model to explain the experimental data, we estimate the regression model, in which the functional form is directly guided by the optimizing behavior in condition (15). Nevertheless, for completeness, and to give an indication of the robustness of our results, we also allow for a non-parametric estimation, where the equality condition $(y - 2x)$ and the equality intention $(\hat{y} - 2x)$ are not specifically assumed any predetermined functional form.

5.1 Parametric Estimation

Condition (15) indicates the optimizing behavior for the mixed type of Actor 2 can be estimated by the following equation:

$$y^* = \frac{-1}{2a_1} + \left[\frac{b_1 + 4a_1}{2a_1}\right]x - \left[\frac{b_2}{2a_1}\right]x^2 - \left[\frac{b_3}{2a_1}\right](\hat{y} - 2x)^2 + \left[\frac{a_2}{2a_1}\right]D + \varepsilon$$

where a_1, a_2, b_1, b_2, b_3 measure the degree to which Actor 2 cares about inequality aversion, altruism, reciprocity, and intention towards equality, respectively. The estimated coefficients

obtained from the regression allow us to identify the values of these parameters and the shape of

the utility function. The results of the regression estimates are summarized in Table 3.

Table 3: Regression Estimates for Actor 2's Back Transfer

Dependent variable: Actor 2's back transfer

Variable	Model [1]		Model [2]	
	Coefficient	Std. Err.	Coefficient	Std. Err.
Constant	-7.081***	1.487	2.950**	1.424
x	3.000***	0.519	2.978***	0.522
x^2	-0.121***	0.043	-0.119***	0.043
$(\hat{y} - 2x)^2$	-0.043***	0.014	-0.043***	0.014
D	9.218***	0.879	-9.171***	0.903
Treatment effects				
T2			-0.651	1.165
T3			-1.809	1.246
T4			-0.819	1.168
R^2	0.65		0.66	
Observation	130		130	

* Significance at 10%
** Significance at 5%
*** Significance at 1%

As shown in Table 3, all of the estimated coefficients reveal to have an expected sign in

both regression models. The coefficients on the transfer variable (3.000 in Model [1] and 2.950 in

Model [2]) are positive and significant ($p < 0.01$), suggesting that trust reinforces trustworthiness.

However, this positive relationship exhibits diminishing returns since the coefficient on the

squared transfer variable (-0.121 in Model [1] and -0.119 in Model [2]) is negative and significant

($p < 0.01$). Further, Model [2] indicates no treatment effects on Actor 2's trustworthiness.

In Model [1], the negative coefficient on the equality intention (-0.043) is statistically

significant ($p < 0.01$). This result indicates Actor 2 cares about Actor 1's intention towards

equality and dislike an unfair proposal intentionally made by Actor 1. The coefficient on

conditional variable (9.218) is positive and significant ($p < 0.01$), indicating the presence of the

altruistic preference in Actor 2's decision making. The positive coefficient particularly indicates the altruistic Actor 2 tends to send more money back, approximately 9 shanks higher than the equality condition. The value of $R^2 = 0.65$ indicates a large percentage of the variation in Actor 2's trustworthy behavior can be rationalized by the combination of these social preferences.

To identify the shape of Actor 2's utility function, we then determine the value of the preference parameters by equalizing the estimated coefficients from the regression to the theoretical prediction in Condition (15). The computed values of these parameters are summarized in Table 4.[11]

Table 4: Estimated Preferences from the Regression

Parameter	value
a_1	0.071
a_2	1.302
b_1	0.141
b_2	0.017
b_3	0.006

Using condition (15), we can re-write the utility function as:

16) $u_2^{TM} = 10 - y + 3x - 0.071(y - 2x)^2 + (0.141x - 0.017x^2)y - 0.006(\hat{y} - 2x)^2 y + 1.302(y - 2x)D$

The utility function in (16) confirms our theoretical prediction that material self-interest is not a single source of motivation for Actor 2 as predicted by standard economic theory. In fact, Actor 2 is motivated to exhibit trustworthiness by multiple types of social preferences.

For an altruist, his utility function can be expressed as $u_2^{TM} = 1.302(y - 2x)D$. This specific form of utility function indicates if Actor 2 can help Actor 1 to achieve a higher payoff in

---

[11] Appendix C: Mathematical note for the estimated preferences.

the game ($D = 1$; $\pi_2 < \pi_1$), he receives the additional utility (of 1.302 units) for each shank of

the back transfer above the equality condition. On the other hand, the utility function of the

inequality-averse Actor 2 $u_2^{TM} = -0.071(y - 2x)^2$ indicates Actor 2 suffers the loss (of 0.071

unit) for each shank of the back transfer above or below the equality condition. More precisely,

the inequality-averse Actor 2 maximizes his utility only when the back transfer is roughly twice

the amount sent by Actor 1. The estimated utility functions for altruism and inequality aversion

are presented graphically in Figure 3.

Figure 3: Altruism and Inequality Aversion.
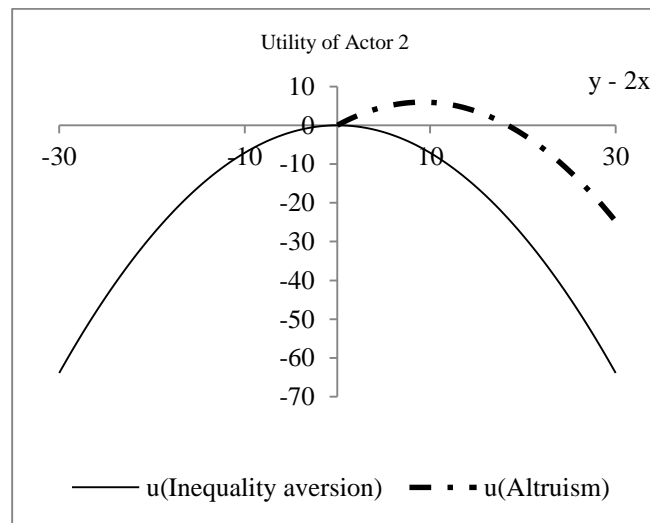
The utility function of Actor 2 who is motivated by the preference for positive reciprocity

can be written as $u_2^{TM} = (0.141x - 0.017x^2)y$. This indicates, for a given amount of the back

transfer, this type of Actor 2 receives additional utility when observing an increase in the transfer

amount sent by Actor 1. However, the diminishing returns (0.017) suggest that the utility from

exhibiting positive reciprocity will reach its maximum level, and eventually start to decline at a

certain level of the transfer. Figure 4 presents the estimated utility function for positive

reciprocity.

Figure 4: Positive Reciprocity
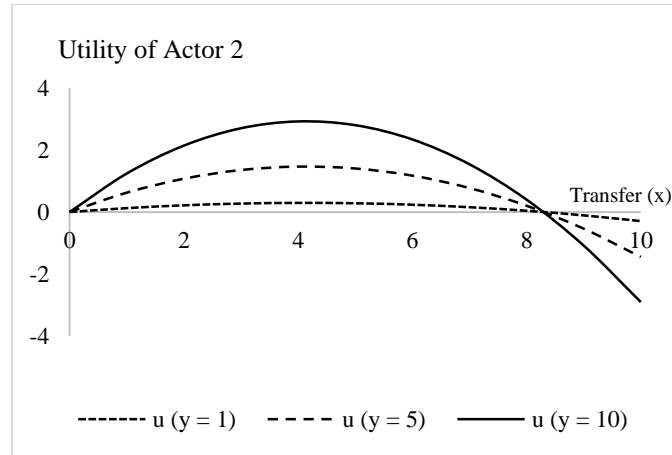
The utility function of Actor 2 who places a strong preference for Actor 1's intention toward equality can be written as $u_2^{TM} = -0.006(\hat{y} - 2x)^2 y$. Since $du_2^{TM}/dy = -0.006(\hat{y} - 2x)^2$, Actor 2's utility is negatively affected by the intended deviation from the equitable outcome proposed by Actor 1, irrespective of whether it is disadvantageous or advantageous inequality. This type of Actor 2 appreciates Actor 1 who intends split all earnings equally by setting $\hat{y} = 2x$, and reward this intention with positive reciprocity. Figure 5 presents the estimated utility function for intention towards equality.

Figure 5: Intention Appreciation
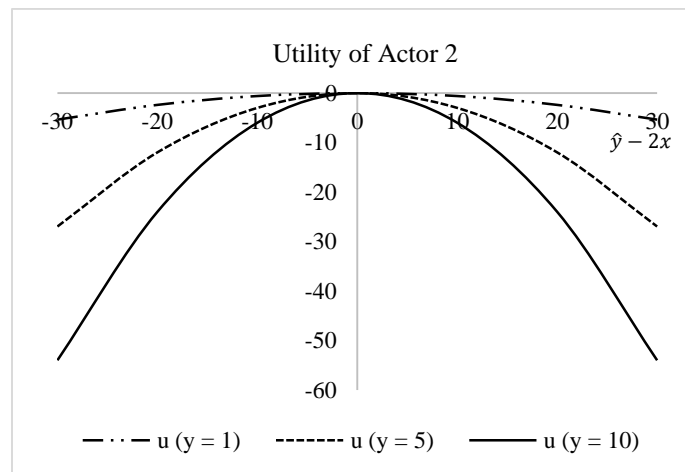
5.2 Robustness Checks

As mentioned, we also perform the non-parametric estimation, where the equality condition $(y - 2x)$ and the equality intention $(\hat{y} - 2x)$ are allowed to take the non-predetermined functional form. We first analyze the nature and distribution of these two variables and categorize them into a set of dummy variables. Figure 6 presents the distribution of each variable, and Table 5 provides a summary of the dummy variables.

Figure 6: Distributions of Equality Condition and Equality Intention
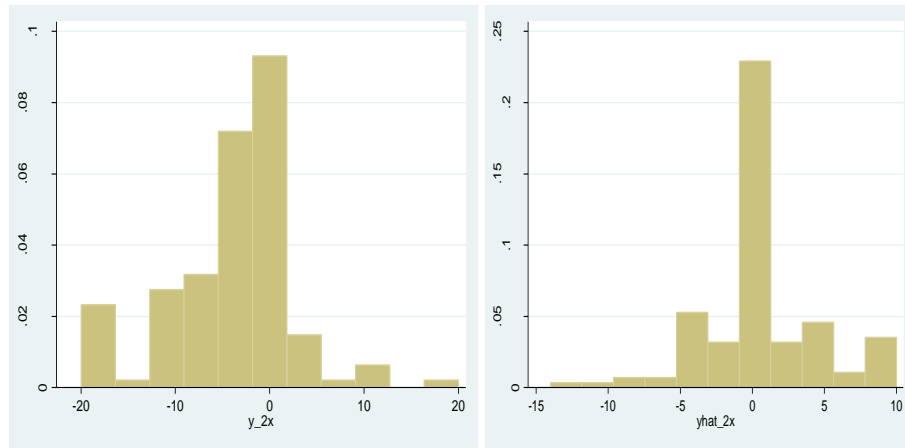
Table 5: A Summary of Dummy Variables

$y - 2x$		$\hat{y} - 2x$	
Dummy variables	Range	Dummy variables	Range
r_1	1 $if\ y - 2x < 0$ 0 $otherwise$	s_1	1 $if\ \hat{y} - 2x \leq -9.2$ 0 $otherwise$
r_2	1 $if\ y - 2x = 0$ 0 $otherwise$	s_2	1 $if\ -9.2 < \hat{y} - 2x \leq -4.4$ 0 $otherwise$
r_3	1 $if\ y - 2x > 0$ 0 $otherwise$	s_3	1 $if\ -4.4 < \hat{y} - 2x \leq 0.4$ 0 $otherwise$
		s_4	1 $if\ 0.4 < \hat{y} - 2x \leq 5.2$ 0 $otherwise$
		s_5	1 $if\ \hat{y} - 2x \leq 5.2$ 0 $otherwise$

Substituting the dummy variables presented in Table 5 into condition (15), we then can rewrite the utility function as

17) $u_2^{TM} = 10 - y + 3x - a_1(y - 2x)^2 + (b_1 x - b_2 x^2)y - (d_1 s_1 + d_2 s_2 + d_3 s_3 +$

$$d_4 s_4 + d_5 s_5)y + (k_1 r_1 + k_2 r_2 + k_3 r_3)(y - 2x)$$

The first-order condition implies

$$\frac{du_2^{TM}}{dy} = -1 - 2a_1(y - 2x) + (b_1 x - b_2 x^2)y - (d_1 s_1 + d_2 s_2 + d_3 s_3 + d_4 s_4 + d_5 s_5)$$

$$+ (k_1 r_1 + k_2 r_2 + k_3 r_3) = 0$$

To maximize the utility, Actor 2 will choose the optimal level of the back transfer at

18) $y^* = \frac{-1}{2a_1} + \left[\frac{b_1 + 4a_1}{2a_1}\right] x - \left[\frac{b_2}{2a_1}\right] x^2 - \left[\frac{d_1}{2a_1}\right] s_1 - \left[\frac{d_2}{2a_1}\right] s_2 - \left[\frac{d_3}{2a_1}\right] s_3 - \left[\frac{d_4}{2a_1}\right] s_4 -$

$$\left[\frac{d_5}{2a_1}\right] s_5 + \left[\frac{c_1}{2a_1}\right] r_1 + \left[\frac{c_2}{2a_1}\right] r_2 + \left[\frac{c_3}{2a_1}\right] r_3$$

Define $\gamma_0 = \frac{-1}{2a_1}$, $\gamma_1 = \frac{b_1 + 4a_1}{2a_1}$, $\gamma_2 = -\frac{b_2}{2a_1}$, $k_1 = \frac{c_1}{2a_1}$, $k_2 = \frac{c_2}{2a_1}$, $k_3 = \frac{c_3}{2a_1}$, $m_1 = -\frac{d_1}{2a_1}$, $m_1 =$

$-\frac{d_2}{2a_1}$, $m_2 = -\frac{d_3}{2a_1}$, $m_3 = -\frac{d_3}{2a_1}$, $m_4 = -\frac{d_4}{2a_1}$, $m_5 = -\frac{d_5}{2a_1}$, then we can rewrite condition (18)

as

19) $y^* = \gamma_0 + \gamma_1 x + \gamma_2 x^2 + m_1 s_1 + m_2 s_2 + m_3 s_3 + m_4 s_4 + m_5 s_5 + k_1 r_1 + k_2 r_2 + k_3 r_3$

Therefore, we can apply the least squares regression to estimate Actor 2's optimizing behavior. The result from the regression is summarized in Table 6.

Table 6: Robustness Test in Opponent's Intention

Dependent variable: Actor 2's back transfer

Variable	Coefficient	Std. Err.
Constant	-6.632***	1.486
x	2.924***	0.512
x^2	-0.111**	0.044
s_1	-5.538*	3.136
s_2	-1.884	1.289
s_4	-1.792	1.167
s_5	-4.592***	1.368
r_2	7.367***	1.009
r_3	13.220***	1.320
Observation	130	
R^2	0.69	

* Significance at 10%
** Significance at 5%
*** Significance at 1%

As can be seen from Table 6, the signs and significant levels of all coefficients are consistent with the parametric regression. Consider the equality intention variable we can see that the negative coefficients on s_1 and s_5 variables are significantly different from zero. This implies the function of back transfer chosen by Actor 2 is strictly concave, and it has a relative maximum in the domain of Actor 1's intention towards equality. In the other words, Actor 2 appreciates the intention toward the equitable outcomes proposed by Actor 1, and reward the equality intention by sending back the desired amount. However, if Actor 1 does not propose equality, Actor 2 experiences a significant loss in his utility, and corrects this inequality intention by sending back less than the desired amount. Figure 7 presents the relationship between Actor 2's optimizing behavior and Actor 1's equality proposal.

Figure 7: Relationship between the Back Transfer and the Equality Intention
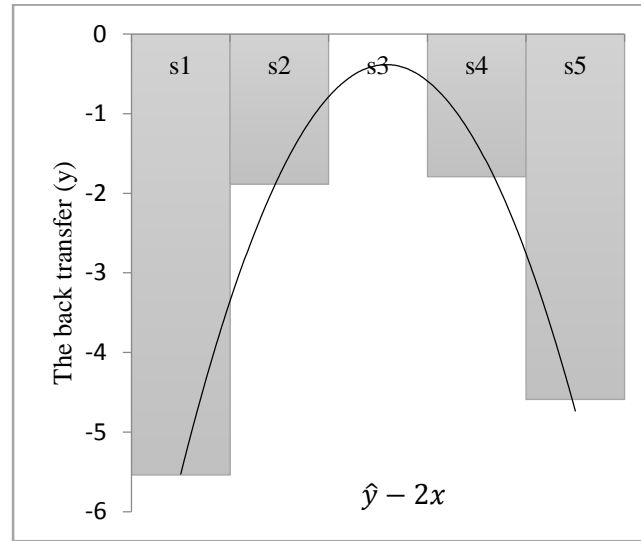
Following the same approach, we estimate the values of the preference parameters in Table 7 by

equalizing the theoretical parameters in condition (18) and the estimated coefficients from the

regression.

Table 7: Estimated Preferences, Robustness Test[12]

Parameter	value
a1	0.075
b1	0.139
b2	0.017
d1	0.835
d2	0.000
d3	0.000
d4	0.000
d5	0.692
c1	0.000
c2	1.111
c3	1.993

---

[12] Appendix D: Mathematical note for the estimated preferences, robustness checks.

Substituting the values presented in Table 8 into condition (18), we can rewrite the utility function as

20) $u_2^{TM} = 10 - y + 3x - 0.075(y - 2x)^2 + (0.139x - 0.017x^2)y - (0.835s_1 + 0.692s_5)y + (1.111r_2 + 1.993r_3)(y - 2x)$

Condition (20) confirm our previous findings that Actor 2 are strongly motivated to exhibit trustworthiness by multiple types of social preferences including inequality aversion, positive reciprocity, altruism, and intention towards equality.

6. Actor 1's Optimizing Behavior

In the following, we illustrate that cooperative equilibria in a one-shot game may exist if Actor 1 has the ability to recognize the preferences and optimizing behavior of Actor 2.

Suppose Actor 1 is motivated solely by self-interest, and anticipates multiples of social preferences in Actor 2's decision making. Following Stackelberg's Model of Duopoly, the optimizing behavior of Actor 1 in the first stage of the game can be analyzed using the concept of backward induction, beginning with the optimizing behavior of Actor 2 in the second stage. Here, we present five possible cases of Actor 1's behavior in choosing the transfer level.[13]

Case 1: Actor 1 anticipates that Actor 2 is motivated purely by self-interest.

Knowing the utility function of the self-interested Actor 2 is the game payoff $u_2 = 10 - y + 3x$, Actor 2 will not provide any back transfer because his marginal utility of exhibiting trustworthiness is strictly negative $du_2/dy - 1 < 0$. Anticipating this, Actor 1 will not transfer anything to Actor 2. Thus, the Nash equilibrium prediction is $(x^* = 0, y^* = 0)$.

---

[13] Appendix E: Mathematical note for Actor 1's optimal behavior.

Case 2: Actor 1 anticipates that Actor 2 is motivated by a composition of self-interest and inequality aversion.

The utility function of Actor 2 who has the preference for self-interest and inequality aversion is given by $u_2 = 10 - y + 3x - a_1(y - 2x)^2.$ To analyze Actor 1's optimal behavior, we first examine the optimizing behavior of Actor 2, which yields $y^* = -1/2a_1 + 2x.$ Since Actor 1 recognizes Actor 2's reaction in the second stage, he anticipates the amount of the back transfer at $y^*.$ Subsequently, Actor 1's problem in the first stage amounts to

21) $Max_x\ u_1 = Max_x\ [10 - x + y^*] = Max_x\ \left[10 - \frac{1}{2a_1} + x\right]$

Note that the utility of Actor 1 is now positive and increasing in the transfer variable. Also note that $\frac{1}{2a_1}$ will converge to zero if Actor 2 is purely inequality averse. Condition (21) implies Actor 1 can be better off by allocating all of his endowment to the Actor 2 who is motivated by a combination of self-interest and inequality aversion. Thus, the optimizing behavior of Actor 1 is to exhibit the highest degree of trust.

Case 3: Actor 1 anticipates Actor 2 is motivated by a composition of self-interest, inequality aversion, and positive reciprocity.

Under the combined preferences for self-interest, inequality aversion, and positive reciprocity, the utility function for Actor 2 is given by $u_2 = 10 - y + 3x - a_1(y - 2x)^2 + (b_1x - b_2x^2)y.$ Thus, Actor 2 maximizes his utility by setting

22) $y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2$

Since Actor 1 can recognize Actor 2's optimal behavior in the second stage, he should anticipate the back transfer at $y^*.$ Thus, Actor 1's optimization problem in the first stage of the game amounts to

23) $Max_x u_1 = 10 - x + y^* = Max_x \left[10 - x + \left[\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2 \right) x - \frac{b_2}{2a_1} x^2 \right] \right]$

Since $du_1/dx = 1 + \frac{b_1}{2a_1} - \frac{b_2}{a_1} x = 0$, the optimal transfer chosen by Actor 1 is then

24) $x^* = \frac{a_1}{b_2} + \frac{b_1}{2b_2} > 0$

In this case, Actor 1's optimal behavior is to transfer the positive amount $(x^* > 0)$.

Case 4: Actor 1 anticipates that Actor 2 is motivated by a omposition of self-interest, inequality aversion, positive reciprocity, and intention towards equality

The utility function of Actor 2 who has the combined preferences for self-interest, inequality aversion, positive reciprocity, and intention towards equality is given by $u_2 = 10 - y + 3x - a_1(y - 2x)^2 + (b_1 x - b_2 x^2)y - b_3(\hat{y} - 2x)^2 y$. Following the same approach, the optimal back transfer chosen by Actor 2 is

25) $y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2 \right) x - \frac{b_2}{2a_1} x^2 - \frac{b_3}{2a_1} (\hat{y} - 2x)^2$

Since Actor 1 can solve Actor 2's problem, he expects to receive the back transfer at y^* in the second stage. Thus, his optimization problem in the first stage is

26) $Max_x u_1 = 10 - x + y^* = Max_x \left[10 - x + \left[\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2 \right) x - \frac{b_2}{2a_1} x^2 - \frac{b_3}{2a_1} (\hat{y} - 2x)^2 \right] \right]$

The first-order condition indicates

$\frac{du_1}{dx} = -1 + \frac{b_1}{2a_1} + 2 - \frac{b_2}{a_1} x - \frac{b_3}{a_1} (\hat{y} - 2x)(-2) = 0$

Subsequently, the optimal transfer chosen by Actor 1 is

27) $x^* = \frac{1}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1} \right]} + \frac{\frac{b_1}{2a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1} \right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1} \right]} \hat{y} > 0$

Condition (27) indicates Actor 1's optimizing behavior is to transfer the positive amount $(x^* > 0)$, and the optimal transfer depends upon the desired back transfer that Actor 1 would like request. A plausible explanation for this condition is that, while recognizing the preference for intention towards equality in Actor 2's motivation, Actor 1 needs to think more carefully about the desired amount he would like request back before choosing the transfer level.

Case 5: *Actor 1 anticipates that Actor 2 is motivated by a composite of self-interest and all types of social preferences, including inequality aversion, positive reciprocity, intention towards equality, and altruism.*

The utility function of Actor 2 is given by $u_2 = 10 - y + 3x - a_1(y - 2x)^2 + a_2(y - 2x)D + (b_1 x - b_2 x^2)y + b_3(\hat{y} - 2x)^2 y$. For an arbitrary choice of Actor 1's transfer, Actor 2's optimal behavior is then

28) $y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2 + a_2 D$

Note that the only difference between condition (27) and (28) is the fixed term $a_2 D$ that represents the preference of altruism in Actor 2's motivation. The interpretation of the fixed term is the altruistic Actor 2 also cares about the well-being of Actor 1, and he would like to make Actor 1 better off by sending more money.

Since Actor 1 can recognize Actor 2's optimal behavior, he anticipates the back transfer at y^* in the second stage. Therefore, Actor 1's optimization problem in the first stage is given by

29) $Max_x u_1 = 10 - x + y^* =$

$$Max_x \left[10 - x + \left[\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2 + a_2 D\right]\right]$$

The first-order condition implies

$$\frac{du_1}{dx} = -1 + \frac{b_1}{2a_1} + 2 - \frac{b_2}{a_1}x - \frac{b_3}{a_1}(\hat{y} - 2x)(-2) = 0$$

Subsequently, the optimal transfer chosen by Actor 1 is

$$30)\ x^* = \frac{1}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]} + \frac{\frac{b_1}{2a_1}}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]}\hat{y}\ >\ 0$$

Similar to the previous case, condition (30) suggests Actor 1's optimal behavior is to transfer the

positive amount to Actor 2, and the optimal transfer level depends upon the desired back transfer.

Anticipating Actor 2's preferences for inequality aversion ($a_1 > 0$), positive reciprocity

($b_2 > 0$), and intention towards equality ($b_3 > 0$), the coefficient on the desired back transfer,

$\frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]}$, is expected to have a positive sign. This implies the transfer of Actor 1 is increasing

function of his own desired back transfer. Substituting the computed values of all preference

parameters from Table 3 into condition (30), we learn

$$31)\ x^* = 3.45 +\ 0.30\hat{y}$$

Condition (31) indicates Actor 1 is likely to exhibit more trust if he expects more trustworthiness

from Actor 2.

Previous experiments have provided strong evidence supporting this positive relationship.

For example, Neaf and Schuup (2009) report a strong positive correlation (Spearman's rank

correlation "$\rho = 0.18$" at $p < 0.01$ level) between the transfer and the desired back transfer.

Bicchieri et al. (2010) transform the desired back transfer into the percentage of the tripled

investment $\hat{y}/3x$, which is referred to as "the expected reciprocity". This variable is then

regressed against the transfer of Actor 1. The results indicate Actor 1 send the additional $0.10

for each percent increase in the expected reciprocity. Similarly, Sapienza et al. (2007) also report

a strong positive correlation between the desired transfer and the transfer amount above 40

percent of the initial endowment.

6.1 Empirical Evidence of Actor 1's Optimal Behavior

In the following, we measure the quality of our theoretical model in explaining Actor 1's behavior by empirically estimating the relationship between the transfer and the desired back transfer using the same set of experimental data.

Since subjects in the experiment are paired anonymously and they do not know the identity of their opponent, we assume that Actor 1 will anticipate all types of preferences (self-interest, inequality aversion, positive reciprocity, intention towards equality, and altruism) in Actor 2's decision making. Thus, Actor 1's optimizing behavior is

$$x^* = \frac{1}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{b_1}{2a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]}\hat{y}$$

We can also write it as

$$x^* = \frac{\left[1 + \frac{b_1}{2a_1}\right]}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]}\hat{y}$$

Define $w_0 = \frac{\left[1+\frac{b_1}{2a_1}\right]}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]}$ and $w_1 = \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1}+\frac{4b_3}{a_1}\right]}$, then we have

$$32) \quad x^* = \omega_0 + \omega_1\hat{y}$$

If Actor 1 anticipates all types of preferences in Actor 2's motives as we hypothesized, the estimated coefficients from the regression should fairly be identical to the theoretical prediction in condition (31). The regression results for Equation 32) are summarized in Table 8.

Table 8: Regression Estimates for Actor 1's Transfer

Dependent variable: Actor 1's transfer

Variables	Model [1]		Model [2]	
	Coef.	Std. Err.	Coef.	Std. Err.
Constant	1.458***	0.322	1.340***	0.474
$\hat{y}$	0.376***	0.021	0.376***	0.021
Treatment effects				
T4			0.097	0.468
T3			0.242	0.510
T4			0.129	0.465
R2	0.72		0.72	
Observation	130		130	

* Significance at 10%

** Significance at 5%

*** Significance at 1%

The positive coefficients on the desired back transfer are significant ($p < 0.01$) in both models, confirming our hypothesis that Actor 1's transfer is exogenously determined by the desired amount he would like request back. That is, Actor 1 will send addition 0.376 shank to Actor 2 for each shank increase in his desired back transfer. The overall fit, $R^2 = 0.72$, seems reasonable for such a diverse group of Actor1 subjects in this cross-sectional data.

However, we observe a slightly difference between the theoretical prediction (0.300) in condition (31) and the estimated coefficients (0.376) in Table 8. A plausible explanation for this difference is that Actor 1 may not be a profit maximizer who is motivated purely by the self-interest, and there may be multiple types of social preferences (e.g., inequality aversion, altruism, and reciprocity) that plays role in his decision making as well. Future research could seek to integrate these social preferences into Actor 1's utility function. This would provide valuable insight into the determinant of trusting behavior.

7. Summary

In summary, this paper makes an important contribution to the literature by developing a theoretical model of play in the trust game, where the preferences of second mover are a

composite of material self-interest and multiple types of social preferences including fairness, positive reciprocity, and altruism. The model specifically features the potential which proposing equality has for promoting positive reciprocal behavior. In particular, the second mover appreciates the first mover that intends to an equitable outcome and rewards this intention with positive reciprocity.

The prediction of our theoretical model is strongly consistent with the evidence collected from the trust experiment. Our empirical results suggest, when the first mover proposes equality by indicating the desired back transfer roughly twice the amount sent, the second mover maximizes his utility and reciprocates this equality intention by sending back the desired amount. On the other hand, when the first mover does not propose equality, the second mover experiences the utility loss from this inequality intention and responses by sending back less than the desired amount.

Knowing the optimizing behavior and preference of the second mover, we then disclose the linkages between trust and trustworthiness by analyzing the first mover's optimizing behavior. We illustrate the ability to recognize the second mover's preferences will encourage the first mover to exhibit more trust and behave more cooperatively. In addition, our model indicates the degree of trust exhibited by the first mover can be explained exogenously by his own desired back transfer, or the expectation on the second mover's trustworthiness. Consistent with the theoretical prediction, our empirical results suggest the first mover will send additional 0.376 shank to the second mover for each shank increase in the desired back transfer. The interpretation here is, once the first mover can recognize the preference for equality intention in the second mover, he think more carefully and more strategically about the desired amount he would like to request back before making the transfer decision. It is an interesting task for future research to examine the validity and reliability of the model with different sets of experimental data.

The model in this paper assumes that the first mover is motivated purely by self-interest, and aim to maximize his own benefit. However, it is not necessarily for the first mover to exhibit trust only because of his ability to recognize the preference of the second mover. There may also be multiple types of social preferences that encourage the trusting behavior (e.g., fairness, altruism, generosity, or reciprocity). It is an interesting task for future research to integrate these preference into the first mover's motives. This would provide valuable insight into the determinant of trust.

8. References:

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy*, 97, 1447–1458.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the rationality of altruism. *Econometrica*, 70, 737–753.

Bacharach, M., Guerra, G., & Zizzo, D. (2007). The self-fulfilling property of trust: An experimental study. *Theory and Decision*, 63(4), 349-388.

Barr, A. (2003). Trust and Expected Trustworthiness: Experimental evidence from Zimbabwean villages. *The Economic Journal*, 113 (489), 614-630.

Ben-Ner, A., Putterman, L., & Ren, T. (2011). Lavish returns on cheap talk: Two-way communication in trust games. *Journal of Socio-Economics*, 40(1), 1–13.

Berg, J., Dickhaut J., & McCabe K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 122-142.

Bicchieri, C., Lev-On, A., & Chavez, A. (2010). The medium or the message?: Communication richness and relevance in trust games. *Synthese*, 176(1), 125-147.

Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., & Pancotto, F. (2013). Cooperation hidden frontiers: The behavioral foundations of the Italian North-South divide. *Economics Working Papers ECO2013/04*, European University Institute.

Bolton, G. E. (1991). A comparative model of bargaining: Theory and evidence. *American Economic Review*, 81, 1096–1136.

Bolton, G., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity and competition. *American Economic Review*, 90, 166–193.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Russell Sage Foundation, New York, Princeton University Press.

Charness, G. & Dufwenberg, M. (2006). Promises and partnership. *Econometrica,* 74, 1579-1601.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117, 817–869.

Cochard, F., Nguyen-Van, P., Willinger, M. (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior & Organization*, 55, 31–44.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260–281.

Cox, J. C., Friedman, D., & Gjerstad, S. (2006). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1), 17–45.

Cox, J. C., Sadiraj, K., & Sadiraj, V. (2001). Trust, fear, reciprocity and altruism. Mimeo. University of Arizona.

Duffy, J., & Feltovich, N. (2002). Do Actions Speak Louder Than Words? An experimental comparison of observation and cheap talk. *Games and Economic Behavior*, 39(1), 1-27.

Engelmann, D. (2012). How not to extend models of inequality Aversion. *Journal of Economic Behavior & Organization,* 81(2), 599-605.

Erlei, M. (2004). Heterogeneous social preferences. Mimeo. Clausthal University of Technology.

Falk, A., Fehr, E., & Fischbacher, U., (2008). Testing theories of fairness: Intentions matter. *Games and Economic Behavior*, 62(1), 287-303.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14, 159–181.

Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives – trust and trustworthiness among CEOs. *Journal of the European Economic Association,* 2(5), 743–771.

Fehr, E., & Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature* 422, 137–140.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics,* 114, 817-51.

Fehr, E. & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism: Experimental evidence and new theories. In Kolm S. and Ythier, J. M. (ed.), 2006. "*Handbook of the Economics of Giving, Altruism and Reciprocity*." Volume 1 Elsevier, New York.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring Trust. *Quarterly Journal of Economics.* 115(3), 811-846.

Huck, S., Müller, W., & Normann, H.-T. (2001). Stackelberg beats cournot: On collusion and efficiency in experimental markets. *Economic Journal*, 111, 749–766.

Isaac, M., & Walker, J. (1998). Nash as an Organizing Principle in the Voluntary Provision of Public Goods: Experimental evidence. *Experimental Economics*, 1(3), 191-206.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32, 865‑889.

Kirchsteiger, G. (1994). The role of envy in ultimatum games. *Journal of Economic Behavior & Organization*, 25, 373–389.

Kohler, S. (2011). Altruism and fairness in experimental decisions. *Journal of Economic Behavior & Organization,* 80(1), 101-109.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1, 593–622.

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

Naef, M., & Schupp, J. (2009). Measuring trust: Experiments and surveys in contrast and combination. *IZA Discussion Papers 4087*, Institute for the Study of Labor (IZA).

Ottone, S., & Ponzano, F. (2005). An extension to the model of inequity aversion by Fehr and Schmidt. *POLIS Working Papers 51,* Institute of Public Policy and Public Choice.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83, 1281–1302.

Rankin, D. J., & Taborsky, M. (2009). Assortment and the evolution of generalized reciprocity. *Evolution, 63,* 1913–1922.

Rotemberg, J. (2004). Minimally acceptable altruism and the ultimatum game. Mimeo. Harvard Business School.

Sapienza, P., Simats, A. T., & Zingales, L. (2007). Understanding trust. *CEPR Discussion Papers 6462*, C.E.P.R. Discussion Papers.

Sears, D. O., & Funk, C. L. (1991). The role of self-interest in social and political attitudes. In M. Zanna (Ed.), *Advances in Experimental Social Psychology*, 24, 1 -91.

Sheremeta, R., & Zhang, J. (2013). Three-player trust game with insider communication. *MPRA Paper 43533*, University Library of Munich, Germany.

Willinger, M., Keser, C., Lohmann, C., & Usunier, J.-C. (2003). A Comparison of Trust and Reciprocity between France and Germany: Experimental investigation based on the investment game. *Journal of Economic Psychology*, 24(4), 447-466.

Xiao, E., & Bicchieri, C. (2010). When equality trumps reciprocity: Evidence from a laboratory experiment. *Journal of Economic Psychology*, 31, 456-470.

Appendix

Appendix A: Mathematical note for theoretical model.

The utility function for mixed types of Actor 2 is given by

$$u_2^{TM} = 10 - y + 3x - a_1(y - 2x)^2 + a_2(y - 2x)D_2 - b_1(\hat{y} - 2x)^2 y + b_2 xy - b_3 x^2 y$$

First-order condition:

$$\frac{du}{dy} = -1 - 2a_1(y - 2x) + a_2 D_2 - b_1(\hat{y} - 2x)^2 + b_2 x - b_3 x^2 = 0$$

$$2a_1(y - 2x) = -1 + a_2 D_2 - b_1(\hat{y} - 2x)^2 + b_2 x - b_3 x^2 = 0$$

Rearrange the terms, we can re-write it as

$$y = = \frac{-1}{2a_1} + \frac{a_2}{2a_1}D_1 - \frac{b_1}{2a_1}(\hat{y} - 2x)^2 + \frac{(b_2 + 4a_1)}{2a_1}x - \frac{b_3}{2a_1}x^2$$

Appendix B: Instructions for the experiment.

<div align="right">**You are actor 1**</div>

Description of Your Decision Problem

You are a participant in the following decision-making problem. You have been randomly matched with another participant in this problem who is in another room. You will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity. You are in the role of actor 1 and the matched participant is in the role of actor 2. You as well as actor 2 participate only once in this decision problem. You make your decisions with the help of the decision sheet that has been handed out together with this description. Here are the rules that you and actor 2 have to obey when you make your decisions:

<u>Endowment</u>

At the beginning both actors receive an initial endowment of 10 shanks (experimental currency units)

<u>Your decision</u>

You have to make a decision that consists of two components:

1) **A *transfer* between 0 and 10 shanks <u>to</u> actor 2.**

> You can transfer any amount between 0 and 10 shanks **to actor 2**. You make this decision by indicating a number between 0 and 10 in the appropriate box on your decision sheet. We will then triple this transferred amount, i.e., actor 2 receives three times the amount of shanks you transferred.

2) **A *desired back-transfer* <u>from</u> actor 2.**

> After you have made your transfer to actor 2 you indicate a desired back-transfer on your decision sheet. The desired back-transfer is the amount you would like to receive back from actor 2. The desired back-transfer can be any number between 0 and three times the amount you have transferred.

<u>The decision of actor 2</u>

Once you have fixed both components of your decision sheet, we collect your decision sheet and give it to actor 2. In this way we inform actor 2 about your decisions. The actor 2 can transfer any amount of the total number of shanks he received **back to you**.

Payoffs

You as actor 1 received: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

Actor 2 receives: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1.

<u>Exchange rate</u>: For every shank you earn you will be paid $2 (2 U.S. dollars).

You are actor 1

Description of a New Decision Problem

You now participate in a new decision-making problem. As before, you have been randomly matched with another participant in another room. You are again in the role of **actor 1**. The other participant is in the role of actor 2. Notice that in this new decision problem you are matched with a **new person**, i.e., actor 2 is now a different person compared too the previous problem. Once again, you will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity.

The new decision problem is—with one exception—identical to the previous problem. The exception concerns the conditional payoff cut. ***In the new problem you can impose a condition payoff cut of 4 shanks on actor 2.*** In every other respect the problem is the same. Thus both actors again receive an initial endowment of 10 shanks.

Your decision

Again you have to indicate on your decision sheet what amount you want to transfer to actor 2 and what your desired back-transfer is. Actor 2 receives three times the amount of shanks you transferred.

In addition to the transfer and desired back-transfer you also have to indicate on your decision sheet if you want to impose a conditional payoff cut of 4 shanks on actor 2.

- A conditional payoff cut of 4 shanks for actor 2 has the following consequences: The payoff of actor 2 will be reduced by 4 shanks if his actual back-transfer is less than your desired back-transfer. The conditional payoff cut is <u>not</u> due, i.e., it does not reduce the income of actor 2, if actor 2 transfers exactly your desired amount or more to you
- If you do not impose a conditional payoff cut—the income of actor 2 will not be reduced, irrespective of how large the back-transfer of actor 2 is.

The decision of actor 2

Once you have fixed all three components of your decision sheet, we collect your decision sheet and give it to actor 2. In this way we inform actor 2 about your decisions. The actor 2 can transfer any amount of the total number of shanks he received **back to you**. In case that you have chose a conditional payoff cut of 4 shanks, and if actor 2 transfers back less than what you desired, the conditional payoff cut is due.

Payoffs

You as actor 1 received: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

Actor 2 receives: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1 - 4 shanks (in case that a conditional payoff cut has been imposed **and** is due).

<u>Exchange rate</u>: For every shank you earn you will be paid $2 (2 U.S. dollars).

Description of Your Decision Problem

You are a participant in the following decision-making problem. You have been randomly matched with another participant in this problem who is in another room. You will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity. You are in the role of **actor 2** and the matched participant is in the role of actor 2. You as well as actor 1 participate only once in this decision problem. You make your decisions with the help of the decisions sheet that has been handed out together with this description. Here are the rules that you and actor 1 have to obey when you make your decisions:

Endowment

At the beginning both actors receive an initial endowment of 10 shanks (experimental currency units).

The decision of actor 1

First actor 1 has to make a decision that consists of the following two components:

1) **A *transfer* between 0 and 10 shanks <u>to</u> actor 2.**

> Actor 1 can transfer any amount between 0 and 10 shanks **to you**. Actor 1 makes this decision by indicating a number between 0 and 10 in the appropriate box on the decision sheet. We will then triple this transferred amount, i.e., you will receive three times the amount of shanks you transferred.

2) **A *desired back-transfer* <u>from</u> actor 2.**

> After actor 1 has made a transfer to you he indicated a desired back-transfer on your decision sheet. The desired back-transfer is the amount he would like to receive back from you. The desired back-transfer can be any number between 0 and three times the amount that actor 1 has transferred to you.

Your decision

Once actor 1 has fixed both components of the decision, we collect the decision sheet and give it to you. In this way we inform you about actor 1's decisions. Then you can transfer any amount of the total number of shanks you received **back to actor 1**.

Payoffs

Actor 1 receives: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

You as actor 2 receive: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1.

<u>Exchange rate</u>: For every shank you earn you will be paid $2 (2 U.S. dollars).

Description of a New Decision Problem

You will now participate in a new decision-making problem. As before, you have been randomly matched with another participant in another room. You are again in the role of **actor 2**. The other participant is in the role of actor 1. Notice that in this new decision problem you are matched with a **new person**, i.e., actor 1 is now a different person compared too the previous problem. Once again, you will never be informed of the identity of this person, either during or after the experiment; similarly, your matched participant will never be informed about your identity.

The new decision problem is—with one exception—identical to the previous problem. The exception concerns the conditional payoff cut. ***In the new problem actor 1 can impose a condition payoff cut of 4 shanks on you.*** In every other respect the problem is the same. Thus both actors again receive an initial endowment of 10 shanks.

The decision of actor 1

Again actor 1 has to indicate on the decision sheet what amount he wants to transfer to you and what his desired back-transfer is. You receive three times the amount of shanks actor 1 transferred to you.

In addition to the transfer and desired back-transfer actor 1 also has to indicate on the decision sheet if you want to impose a conditional payoff cut of 4 shanks on you.

- A conditional payoff cut of 4 shanks has the following consequences for you: Your payoff will be reduced by 4 shanks if your actual back-transfer is less than the desired back-transfer of actor 1. The conditional payoff cut is <u>not</u> due, i.e., it does not reduce your income, if you transfer exactly the desired amount or more to actor 1.
- If actor 1 does not impose a conditional payoff cut—your income will not be reduced, irrespective of how large your back-transfer to actor 1 is.

Your decision

Once actor 1 has fixed all three components of the decision, we collect the decision sheet and give it to you. In this way we inform you about actor 1's decisions. Then you can transfer any amount of the total number of shanks received **back to actor 1**. In case that actor 1 imposed a conditional payoff cut of 4 shanks, and if you transfer back less than actor 1's desired amount, the conditional payoff cut is due.

Payoffs

Actor 1 receives: 10 shanks – transfer to actor 2 + back-transfer from actor 2.

You as actor 2 receive: 10 shanks + 3*transfer from actor 1 – back-transfer to actor 1 - 4 shanks (in case that a conditional payoff cut has been imposed **and** is due).

<u>Exchange rate:</u> For every shank you earn you will be paid $2 (2 U.S. dollars).

Appendix C: Mathematical note for the estimated preferences.

$$y = \frac{-1}{2a_1} + \left[\frac{b_1 + 4a_1}{2a_1}\right] x - \left[\frac{b_2}{2a_1}\right] x^2 - \left[\frac{b_3}{2a_1}\right] (\hat{y} - 2x)^2 + \left[\frac{a_2}{2a_1}\right] D$$

Variable	Coefficient	Std. Err.	95% Confident Interval	
Constant	-7.081***	1.487	-10.023	-4.139
x	3.000***	0.519	1.973	4.027
x^2	-0.121***	0.043	-0.206	-0.036
$(\hat{y} - 2x)^2$	-0.043***	0.014	-0.071	-0.015
D	9.218***	0.879	7.478	10.958
Observation	130			
R^2	0.65			

* Significance at 10%

** Significance at 5%

*** Significance at 1%

$$\frac{-1}{2a_1} = -7.08 \rightarrow a_1 = 0.07$$

$$\frac{a_2}{2a_1} = 9.21 \rightarrow a_2 = 1.302$$

$$\frac{(b_1 + 4a_1)}{2a_1} = 3.00 \rightarrow b_1 = = 0.141$$

$$-\frac{b_2}{2a_1} = -0.121 \rightarrow b_2 = 0.017$$

$$-\frac{b_3}{2a_1} = -0.043 \rightarrow b_3 = 0.006$$

Appendix D: Mathematical note for the estimated preferences, robustness checks.

$$y = \frac{-1}{2a_1} + \left[\frac{b_1 + 4a_1}{2a_1}\right]x - \left[\frac{b_2}{2a_1}\right]x^2 - \left[\frac{d_1}{2a_1}\right]s_1 - \left[\frac{d_2}{2a_1}\right]s_2 - \left[\frac{d_3}{2a_1}\right]s_3 - \left[\frac{d_4}{2a_1}\right]s_4 - \left[\frac{d_5}{2a_1}\right]s_5$$
$$+ \left[\frac{c_1}{2a_1}\right]r_1 + \left[\frac{c_2}{2a_1}\right]r_2 + \left[\frac{c_3}{2a_1}\right]r_3$$

Variable	Coefficient	Std. Err.	95% Confident Interval	
Constant	-6.632***	1.486	-9.573	-3.690
x	2.924***	0.512	1.910	3.938
x^2	-0.111**	0.044	-0.197	-0.025
s_1	-5.538*	3.136	-11.747	0.670
s_2	-1.884	1.289	-4.436	0.668
s_4	-1.792	1.167	-4.102	0.518
s_5	-4.592***	1.368	-7.299	-1.884
r_2	7.367***	1.009	5.369	9.364
r_3	13.220***	1.320	10.606	15.834
Observation	130			
R^2	0.69			

* Significance at 10%
** Significance at 5%
*** Significance at 1%

$$\frac{-1}{2a_1} = -6.632 \rightarrow a_1 = 0.075$$

$$\frac{(b_1 + 4a_1)}{2a_1} = 2.924 \rightarrow b_1 = 0.139$$

$$-\frac{b_2}{2a_1} = -0.111 \rightarrow b_2 = 0.017$$

$$\frac{d_1}{2a_1} = -5.538 \rightarrow d_1 = -0.835$$

$$\frac{d_2}{2a_1} = 0 \rightarrow d_2 = 0$$

$$\frac{d_4}{2a_1} = 0 \rightarrow d_4 = 0$$

$$\frac{d_5}{2a_1} = -4.592 \rightarrow d_5 = -0.692$$

$$\frac{c_2}{2a_1} = 7.367 \rightarrow c_2 = 1.111$$

$$\frac{c_3}{2a_1} = 16.220 \rightarrow c_2 = 1.993$$

Appendix E: Actor 1's optimal behavior.

Case 1: Actor 1 believes that Actor 2 is motivated solely by self-interest

$$u_2 = 10 - y + 3x$$

$$\frac{du_2}{dy} = -1 < 0$$

Actor 2 will send back nothing, and Actor 1 anticipates this reaction. Thus, there is no trust.

Case 2: Actor 1 believes that Actor 2 is motivated by a composite of self-interest and the preferences for inequality aversion

$$u_2 = 10 - y + 3x + -a_1(y - 2x)^2$$

$$\frac{du_2}{dy} = -1 - 2a_1(y - 2x) = 0$$

$$2a_1(y - 2x) = -1$$

$$2a_1y = -1 + 4a_1x$$

$$y^* = -\frac{1}{2a_1} + 2x$$

Actor 1:

$$u_1 = 10 - x + y^*$$

$$u_1 = 10 - x + \left[-\frac{1}{2a_1} + 2x\right]$$

$$u_1 = 10 - \frac{1}{2a_1} + x$$

$$\frac{du_1}{dx} = 1 > 0$$

Consequently, Actor 1 is willing to transfer positive amount so as to maximize his utility.

Case 3: Actor 1 believes that there are three types of preferences that motivate Actor 2, including self-interest, inequality aversion, and positive reciprocity

$$u_2 = 10 - y + 3x - a_1(y - 2x)^2 + (b_1x - b_2x^2)y$$

$$\frac{du_2}{dy} = -1 - 2a_1(y - 2x) + b_1x - b_2x^2 = 0$$

$$2a_1(y - 2x) = -1 + b_1x - b_2x^2$$

$$2a_1y = -1 + (b_1 + 4a_1)x - b_2x^2$$

$$y = \frac{-1}{2a_1} + \frac{(b_1 + 4a_1)}{2a_1}x - \frac{b_2}{2a_1}x^2$$

$$y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2$$

Actor 1:

$$u_1 = 10 - x + y^*$$

$$u_1 = 10 - x + \left[\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2\right]$$

$$\frac{du_1}{dx} = -1 + \frac{b_1}{2a_1} + 2 - \frac{b_2}{a_1}x = 0$$

$$\frac{b_2}{a_1}x = 1 + \frac{b_1}{2a_1}$$

$$x^* = \frac{a_1}{b_2} + \frac{b_1}{2b_2} > 0$$

Therefore, the optimal behavior for Actor 1 is to transfer the positive amount.

Case 4: Actor 1 believes that there are four types of preferences that motivate Actor 2 including self-interest, inequality aversion, positive reciprocity, and intention based reciprocity

$$u_2 = 10 - y + 3x - a_1(y - 2x)^2 + (b_1x - b_2x^2)y - b_3(\hat{y} - 2x)^2y$$

$$\frac{du_2}{dy} = -1 - 2a_1(y - 2x) + b_1x - b_2x^2 - b_3(\hat{y} - 2x)^2 = 0$$

$$2a_1(y - 2x) = -1 + b_1x - b_2x^2 - b_3(\hat{y} - 2x)^2$$

$$2a_1y = -1 + (b_1 + 4a_1)x - b_2x^2 - b_3(\hat{y} - 2x)^2$$

$$y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2$$

Actor 1:

$$u_1 = 10 - x + y^*$$

$$u_1 = 10 - x + \left[\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2\right]$$

$$\frac{du_1}{dx} = -1 + \frac{b_1}{2a_1} + 2 - \frac{b_2}{a_1}x - \frac{b_3}{a_1}(\hat{y} - 2x)(-2) = 0$$

$$\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right] x = 1 + \frac{b_1}{2a_1} + \frac{2b_3}{a_1}\hat{y}$$

$$x^* = \frac{1}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{b_1}{2a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]}\hat{y} > 0$$

Therefore, the optimal behavior for Actor 1 is to transfer positive amount to Actor 2, and the amount of transfer is dependent upon the desired back transfer

Case 5: Actor 1 believes that Actor 2 is motivated by a composite of self-interest and all types of social preferences, including inequality aversion, positive reciprocity, and intention based reciprocity, and altruism

$$u_2 = 10 - y + 3x - a_1(y - 2x)^2 + a_2(y - 2x)D + (b_1x - b_2x^2)y + b_3(\hat{y} - 2x)^2 y$$

$$y^* = \frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2 + a_2D$$

Also note that D =1 if $y - 2x > 0$. Altruistic preferences is an additional fixed term on the optimal level of back transfer chosen by Actor 2.

Actor 1:

$$u_1 = 10 - x + y^*$$

$$u_1 = 10 - x + [\frac{-1}{2a_1} + \left(\frac{b_1}{2a_1} + 2\right)x - \frac{b_2}{2a_1}x^2 - \frac{b_3}{2a_1}(\hat{y} - 2x)^2 + a_2D]$$

$$\frac{du_1}{dx} = -1 + \frac{b_1}{2a_1} + 2 - \frac{b_2}{a_1}x - \frac{b_3}{a_1}(\hat{y} - 2x)(-2) = 0$$

$$\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right] x = 1 + \frac{b_1}{2a_1} + \frac{2b_3}{a_1}\hat{y}$$

$$x^* = \frac{1}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{b_1}{2a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]} + \frac{\frac{2b_3}{a_1}}{\left[\frac{b_2}{a_1} + \frac{4b_3}{a_1}\right]}\hat{y} > 0$$

Similar to Case 3, the optimal behavior for Actor 1 is to transfer positive amount to Actor 2, and the amount of transfer is dependent upon the desired back transfer.