Title
Essays on Behavior in Games

Author
Ridinger, Garret

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Essays on Behavior in Games

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Garret Ridinger

Dissertation Committee:
Professor Michael T. McBride, Chair
Professor Stergios Skaperdas
Professor John Duffy
Associate Professor Igor Kopylov

2016

DEDICATION

I would like to dedicate this dissertation to Anne Carpenter, Cindy Hollingworth, and Mark Hollingworth.

My wife, Anne, has been with me from the beginning and I could not have done this without her. Her support and encouragement throughout this process is what made completing this dissertation possible. Anne's intelligent advice helped me grow as both a researcher and as a person. Her unwavering confidence in me has been an inspiration. I am incredibly lucky to know her and want to thank her for everything she has done for me.

My parent's, Cindy and Mark, deserve enormous credit for the person I am today. Their sacrifices and guidance have allowed me to achieve my goals and I am truly indebted to them. I want to thank them for teaching me the value of hard work and never giving up on your dreams.

I would also like to thank my entire family for all their help, support, and understanding. For giving me a much needed break, I would like to thank my brother, Alec, for all the basketball games and my sister, Nicole, for all the laughs.

Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

ACKNOWLEDGMENTS

I would like to thank Anne Carpenter for all her advice, comments, and suggestions which have improved my work tremendously. She has helped me in numerous areas including helping me think through ideas, improving my presentations, and developing my writing skills. I have learned a great deal from her and look forward to working with her in the future.

The chair of my committee, Michael T. McBride, deserves an enormous amount of credit for helping me improve in all aspects of my work. I am forever grateful to him for giving me the opportunity to learn from him and work as his research assistant. I want to thank him for investing an incredible amount of time and energy in helping me be successful. It has been a pleasure working with him and I look forward to our future projects.

I have learned immensely from my committee member, Stergios Skaperdas. I have really enjoyed talking to him about a wide range of topics from new research ideas to current political issues to everyday life. I want to thank him for all his help and look forward to working with him in the future.

I would like to also thank my committee members, John Duffy and Igor Kopylov. Both John and Igor have been a great sources of knowledge and advice. I want to thank them both for all their help and look forward to working with them in the future.

I acknowledge and appreciate financial support from the Army Research Office Award No. W911NF-11-1-0332, the Experimental Social Science Laboratory (ESSL), and the Economics Department at the University of California, Irvine.

CURRICULUM VITAE

Garret Ridinger

EDUCATION

Doctor of Philosophy in Economics 2016
University of California, Irvine

Master of Arts in Economics 2011
California State University, Fullerton

Bachelor of Arts in Business Economics 2007
University of California, Santa Barbara

ACADEMIC POSITIONS

Graduate Research Assistant for Michael McBride 2013-2016
University of California, Irvine

Laboratory Manager for the Experimental Social Science Laboratory 2015
University of California, Irvine

Teaching Assistant 2011-2015
University of California, Irvine

PUBLICATIONS

Ridinger, G., McBride, M. Money Affects Theory of Mind Differently by Gender. 2015
PLOS ONE 10(12): e0143973

Ridinger, G., McBride, M. Attacker Deterrence and Perceived Risk in a StackelbergForthcoming
Security Game. *Risk Analysis*

CONFERENCE PRESENTATIONS

Southwest Experimental and Behavioral Economics Conference 2016
University of California, Irvine

Experimental Social Science Workshop 2015
University of California, Irvine

Bay Area Behavioral and Experimental Economics Workshop. 2015
University of California, Santa Cruz

Theory, History, and Development Seminar Series (2) 2015
University of California, Irvine

Tuesday Talk Seminar Series. 2015
Claremont Graduate University

ESA North American Conference. 2015
Dallas

Brown Bag Workshop at the Economic Science Institute. 2014
Chapman University

Experimental Social Science Graduate Student Workshop. 2014
University of California, Irvine

Bay Area Behavioral and Experimental Economics Workshop. 2014
Stanford University

ESA International Meetings. 2014
Hawaii

SABE Annual Conference. 2014
Lake Tahoe

ESA North American Conference. 2014
Ft. Lauderdale

11th Annual IMBS Graduate Student Conference. 2013
University of California, Irvine

Experimental Social Science Laboratory Experimental Workshop. 2013
University of California, Irvine

ESA North American Conference. 2013
Santa Cruz

ABSTRACT OF THE DISSERTATION

Essays on Behavior in Games

By

Garret Ridinger

Doctor of Philosophy in Economics

University of California, Irvine, 2016

Professor Michael T. McBride, Chair

This dissertation consists of three chapters researching how individuals behave in game-theoretic situations. Chapter 1 investigates how different fairness concerns impact individual decisions to cooperate. I introduce a new version of the sequential prisoners dilemma where there is a chance the first movers choice is reversed. Experimental results show that second mover cooperation is higher when the first mover has little control over their choice and when the second mover is not told what the first mover chose. While subject behavior is consistent with concerns for fairness, the results indicate that these concerns work in ways not predicted by current theoretical models. Chapter 2 studies how prior ownership can influence bargaining over a jointly produced surplus. Experiments varied whether the proposer or responder received the surplus prior to bargaining and the strength of punishment. The results suggest that proposers respect prior ownership when the responder has a strong ability to punish, but not when punishment is weak. Responders respect prior ownership when their ability to punish is weak, but reject at high rates when they have strong punishment. I show that an independent measure of rule following can explain the results suggesting that individual behavior in bargaining is driven in part by adherence to social norms. Chapter 3 explores how guilt, shame, and theory of mind (ToM) influence adherence to social norms. Using psychometric measures, I explore how guilt, shame, and ToM are related to following prior ownership norms in a bargaining experiment. Guilt was not predictive of behavior; however, both shame and ToM were. Individuals who had greater feelings of shame were much more likely

to respect prior ownership of responders and punish proposers who transgressed prior ownership of responders. Individuals who scored higher in ToM were much more likely to respect prior ownership by proposer. These results suggest that individual differences in shame and ToM are important in understanding adherence to bargaining norms.

Chapter 1

Intentions versus outcomes: cooperation and fairness in the sequential prisoner's dilemma with nature

1.1 Introduction

Experimental evidence indicates that people often deviate from maximizing their own monetary payoff. In the one-shot sequential prisoner's dilemma, if each player maximizes her own payoff, then the equilibrium prediction is that both players defect. Despite this, cooperation rates by both players are significant (Bolle and Ockenfels, 1990; Clark and Sefton, 2001; Ahn et al., 2007). To explain cooperation in one-shot games, researchers have suggested that people care about fairness and have incorporated these concerns into game-theoretic models. These theoretical models of fairness can be separated into two types: outcome based and intention based.[1] Outcome-based models

[1] Examples of outcome-based models include Bolton and Okenfels (2000) and Fehr and Schmidt (1999), and intention-based models include Rabin (1993), and Dufwenberg and Kirchsteiger (2004). Models that combine concerns for intentions and outcomes include Levine (1998), Charness and Rabin (2002),Falk and Fischbacher (2006), and Cox et al. (2007).

capture concerns over distributions. An example of an outcome-based model is inequity aversion (Fehr and Schmidt, 1999) which allows people to compare their payoff with others and prefer payoffs that are more equal. Intention-based models allow beliefs about others actions to influence fairness concerns. An example of an intention-based model is the reciprocity model (Dufwenberg and Kirchsteiger, 2004), which captures that people may prefer to be kind to people who are kind to them and punish people who are unkind to them. Both modeling approaches incorporate fairness concerns that are likely important in a wide range of human behavior, but in certain environments the predictions by the two approaches can be quite different. The distinction between intentions and outcomes has real-world applications. For example, in the United States legal code, there are different consequences for being charged with involuntary manslaughter compared to first-degree murder. While the outcome is the same in both cases, the intention behind the homicide matters. Despite it's importance, it is still not fully understood how the the relative strength of individual concerns about intentions and outcomes influences human behavior.

Conditional cooperation in the sequential prisoner's dilemma is consistent with concerns for outcomes (Fehr and Schmidt, 1999) and intentions (Dufwenberg and Kirchsteiger, 2004). Due to this, prior research has been unable to disentangle the two effects to understand their importance in explaining cooperation (Bolle and Ockenfels, 1990; Clark and Sefton, 2001; Ahn et al., 2007). This paper introduces a novel game called the sequential prisoner's dilemma with nature. In the game, the first mover decides whether to cooperate or defect. After the first mover's choice, there is a chance the choice is reversed by nature. After observing both what the first mover chose and the results from nature, the second mover can choose to cooperate or defect. This creates a situation where the first mover may intend to cooperate but due to chance they end up defecting. Since the second mover observes what the first mover intended to do and the outcomes are kept the same, the game can differentiate between the two fairness approaches. The game captures environments in which there is an imperfect correlation between actions and the results of those actions. Therefore, it can shed light on a wide range situations including principle agent problems. For example, an employee can choose to work hard (cooperate) or not work hard (defect) on a project. Hard work does not guarantee that the project will be profitable for the employer but it could make it more likely that the project is a success. After observing the effort level and whether the project was successful, the employer may choose whether to reward the employee (cooperate) or not reward the employee (defect).

Often the intentions of others are not fully observable. To explore how information influences cooperation, this paper introduces a variant of the sequential prisoner's dilemma with nature where the second mover is not told what the first mover chose, but does know the results of nature. This feature captures situations where the second mover must infer the first mover's intended choice based on the first mover's control over their choice and the results of nature. For example, in principal agent problems the employees effort level is often not observable. Instead, the employer only observes whether the project was successful. The employer must infer the effort level that the employee contributed based on the correlation between effort and the success of the project. Comparing the two games can add to our understanding of how information about the person's intended choice influences individual behavior.

This paper examines the relative influence of intentions and outcomes on cooperation. Specifically, I ask: What do existing fairness models predict as information and control changes in the sequential prisoner's dilemma with nature? How well do the models capture what people actually do? To address these questions, I begin theoretically by analyzing the equilibrium predictions of outcome-based, intention-based, and combined fairness models. The modeling approaches predict different equilibrium behavior depending on individual types of players and their preferences. I test the theoretical predictions empirically using a laboratory experiment. The design of the experiment allows the separation of intentions and outcomes as well as tests the role of information.

Theoretical results under perfect information show that the outcome-based model of inequity aversion (Fehr and Schmidt, 1999) predicts that second mover cooperation depends only on the results of nature. Suggesting that cooperation will be unaffected by the first mover's choice. The intention-based model of reciprocity (Dufwenberg and Kirchsteiger, 2004) predicts that cooperation depends only on the first mover's choice and that the results of nature will not affect equilibrium behavior. Specifically, inequity aversion suggests that changes in information or control will not influence equilibrium behavior. However, reciprocity predicts that second mover cooperation will be higher when control by the first mover increases, and cooperation should increase when there is imperfect information about the first mover's choice. To account for the possibility that individuals may care about both outcomes and intentions, I introduce the mixed-concerns model. Using psychological game theory, the mixed-concerns model combines both inequity aversion (Fehr and Schmidt, 1999) and reciprocity (Dufwenberg and Kirchsteiger, 2004) into a single framework. The model allows for heterogeneity in subject's weight of two concerns. If individuals care about both reciprocity and inequity aversion, then there exists an additional equilibrium depending on the relative strength of the two concerns.

After examining the theoretical models, I test the predictions experimentally by varying the chance the first mover's choice is reversed and whether the second mover observes the first mover's choice. The results show that second mover cooperation is higher when the first mover has little control over their choice and when the second mover is not told what the first mover chose. While subject behavior is consistent with concerns for both intentions and outcomes, the results indicate that these concerns work in ways not predicted by current theoretical models. Specifically, conditional cooperation by second movers was higher when control was low. This result is puzzling as it is opposite of what is predicted by models of reciprocity. Using psychometric measures, I find that differences in perspective taking ability provide a potential explanation for the puzzle. In addition, higher empathic concern is found to be correlated with increased conditional cooperation by second movers.

Previous research on the sequential prisoner's dilemma has found that second movers are more likely to cooperate if the first mover cooperates (Bolle and Ockenfels, 1990; Clark and Sefton, 2001; Ahn et al., 2007). This finding is in line with other evidence that conditional cooperation is an important explanation for behavior in social dilemmas (Fischbacher et al., 2001; Chaudhuri and

Paichayontvijit, 2006; Herrmann and Thoni, 2009; Rustagi et al., 2010; Chaudhuri, 2011). Using the sequential prisoner's dilemma, Dhaene and Bouckaert (2010) find evidence that conditional cooperation by second movers matches the theoretical predictions from the reciprocity model of (Dufwenberg and Kirchsteiger, 2004) while Blanco et al. (2011) show that individual measures of inequity aversion (Fehr and Schmidt, 1999) can predict second mover behavior. As a result, it is still unclear whether second mover conditional cooperation in the sequential prisoner's dilemma is due to intention-based reciprocity or outcome-based concerns.

In many game-theoretic situations, intention-based models and outcome-based models give similar predictions. This can make it difficult to examine which concerns may have lead to the observed experimental behavior. One approach to has been to vary alternative choices players could have chosen. Results from Falk et al. (2003), Bolton and Okenfels (2005), and Falk and Kosfeld (2006) suggest that changes in the the alternatives available may have influenced behavior. However, changes in the alternatives seemed to have little or no effect in Stanca (2010) and Charness and Rabin (2002). Another approach compares a treatment were a subject has full control over their choice to a treatment where the subject has no control over their choice. Typically, the subjects choice in the no control treatment is selected via random device. Using this approach, Charness (2004) and Falk et al. (2008) found that intentions were important in explaining subject behavior. While Bolton et al. (1998) found that only outcomes mattered. One potential issue in these experiments is that if individuals feel fundamentally different towards random devices compared to when a person is making a choice, then there could be a confounding variable that may bias the results. To control for this, I keep the random device in all the treatments. What varies is the chance the first mover's choice is reversed. In addition to controlling for potential bias, this feature creates a more realistic situation where people have more or less control, but their intended choices still matter.

The experiment conducted by Charness and Levine (2007) used both a random device and varied the alternatives available. Using a modified gift exchange game, the experiment included a coin flip that determined whether the wage of the employee would be higher or lower than what the employer chose. The potential payoffs for the employee were the same, whether the employer chose a high wage and chance made it lower or the employer chose a low wage and chance made it higher. The results suggest that the intentions of the employer influenced the wage choices by employees. Charness and Levine (2007) did not keep all the potential outcomes constant. Instead the alternatives that could be reached were either very beneficial for the employer or very beneficial for the employee. This is how an employer's choice was viewed as having good intentions or not. One key difference in this paper, is that in the experiment the potential end node payoffs are kept constant irrespective of the first movers choice. The first mover can only make certain outcomes more likely to occur.

Information about what the first mover chose can be potentially important in understanding fairness. Charness and Levine (2007) did not include a treatment where workers did not know what wage the firm selected. While this is realistic in the context of their experiment, generalizing the results to other domains becomes difficult in situations were the first mover's choice is not observable. For example, using a trust game Cox and Deck (2006) examined second mover behavior when the first mover's choice had a 25% chance to be reversed, but the second mover was not told what the first mover chose. The results suggest that second movers gave the first movers the benefit of the doubt. Cox and Deck (2006) did not include a treatment where the first mover's choice was known to the second mover. In this paper, I include a condition where the second mover is told what the first mover chose and a condition where the second mover is not told that information. Potential changes in behavior between these two treatments can shed light on the importance of observing others' intentions.

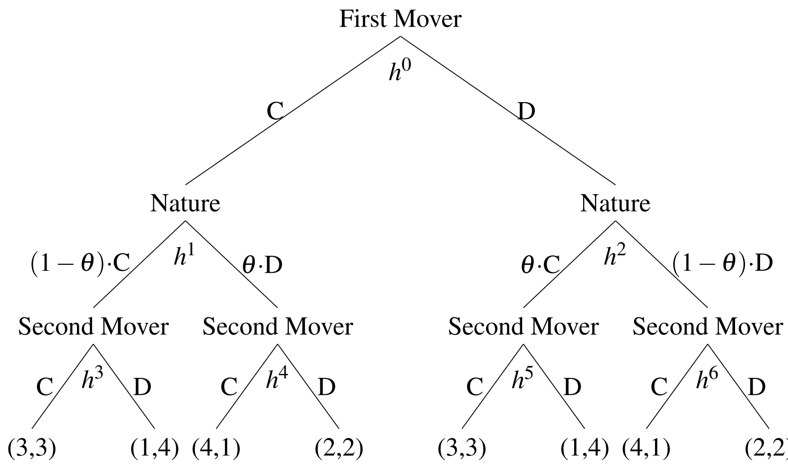


Figure 1.1: Sequential Prisoner's Dilemma with Nature and Perfect Information

1.2 Sequential Prisoner's Dilemma with Nature

Figure 1.1 shows the sequential prisoner's dilemma with nature under perfect information. The first mover decides whether to cooperate or defect first. After first mover chooses, nature will randomly select cooperate or defect. After observing both what first mover chose and the choice by nature, the second mover can choose to cooperate or defect. Figure 1.2 shows the game with imperfect information. In this game, the choice by first mover is no longer observed by the second mover. The second mover only observes whether nature has cooperated or defected.

The probabilities that nature will choose cooperate or defect differ depending on the choice of player 1. The term θ is the chance that the first mover's choice is reversed.[2] When $\theta < \frac{1}{2}$ and the first mover cooperates, there is a higher chance that nature will cooperate compared to if the first mover choose to defect. Natures choice can be thought of as the first mover's control. Lower values of θ make it more likely that that nature will cooperate if the first mover cooperates and defect if the first mover defects. This paper will focus on the case where $\theta \leq \frac{1}{2}$.[3]

[2]While this paper assumes the reversal probability is the same whether the first mover cooperates or defects. Theoretical results are similar if the reversal probabilities are allowed to differ. For clarity of presentation, a single probability θ is used both in the theoretical analysis and the experiment.

[3]The analysis can be done without this restriction. However, if $\theta > \frac{1}{2}$, then the choice that the first mover chooses is more likely to be switched. While this makes sense mathematically, it is not clear that this represents what occurs

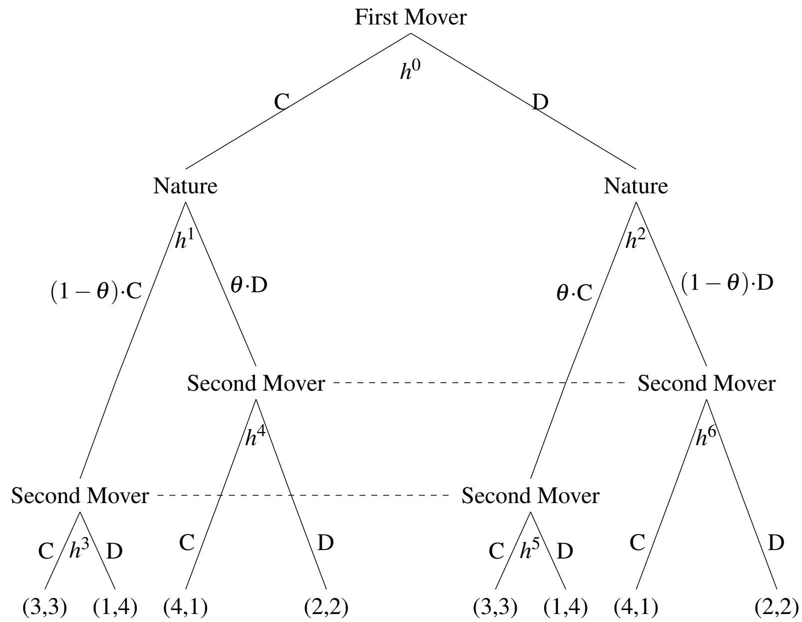


Figure 1.2: Sequential Prisoner's Dilemma with Nature and Imperfect Information

1.3 Theories of Social Preferences

This section provides a review of the theoretical predictions of outcome-based and intention-based models of fairness. For interested readers, formal definitions and equilibrium predictions for the discussed models can be found in the Appendix A. The analysis focuses on second mover behavior, but first mover equilibrium predictions are available upon request.

Proposition 1. *Inequity aversion (Fehr and Schmidt, 1999) predicts that second mover cooperation only depends on the results of nature and information about the first mover's choice will have no effect on cooperate rates*

A prominent model of outcome based preferences is the model of inequity aversion (Fehr and Schmidt, 1999). In a two player game where players have inequity averse preferences (Fehr and

in most human interactions. Having $\theta > \frac{1}{2}$ means that if the first mover wants nature to be more likely to choose cooperate, then the first mover should defect. This is not say these types of situations cannot occur, but the main focus of the paper will be when a player's intended choice matches the player's actual choice.

Schmidt, 1999), each individual i has the following utility function:

$$U_i(\pi_i, \pi_j) = \pi_i - \alpha_i \cdot \max\{\pi_j - \pi_i, 0\} - \beta_i \cdot \max\{\pi_i - \pi_j, 0\} \tag{1.1}$$

where $\alpha_i, \beta_i \geq 0$ and the payoff for individual i is π_i and the payoff for individual j is π_j. Both α_i and β_i capture the degree in which individuals dislike inequality that is advantageous and disadvantageous, respectively. The model captures the idea that people prefer distributions that are more equal.

With perfect information, inequity aversion predicts that cooperation by second movers will only occur if nature cooperates (see proposition 3 in Appendix A for details). In addition, second mover cooperation should not differ depending on whether the first mover cooperated or defected. For second movers, the key parameter that can lead to cooperation is β_i. In order for the second mover to cooperate they must sufficiently dislike getting more than the first mover.

Under imperfect information, inequity aversion predictions are the same. This occurs because the second mover is only concerned about the distribution of outcomes at the end node of the game. As a result, the decision to cooperate is based entirely on the end node payoffs and not how these payoffs were reached. This suggests that changes in information should have no impact on cooperation.

Proposition 2. *Intention-based reciprocity (Dufwenberg and Kirchsteiger, 2004) predicts: (a) that second mover cooperation only depends on the first mover's choice. (b) With perfect information, conditional cooperation is only possible when the reversal probability is low. (c) Cooperation should be higher when information is imperfect.*

According to the intention-based model of reciprocity by Dufwenberg and Kirchsteiger (2004) the utility of an individual i is as follows: $U_i = \pi_i + \lambda_i \cdot k_{ij} \cdot \phi_{iji}$ where $\lambda_i \geq 0$ and captures i's sensitivity towards reciprocity. The function k_{ij} captures person i's kindness towards person j. While the function ϕ_{iji} is a measure of i's belief about the kindness of j towards i. Both k_{ij} and ϕ_{iji} depend on individual first and second order beliefs (for details see Appendix A). This framework measures kindness at a particular node based on the difference between the resulting payoff and an equitable payoff. The equitable payoff is computed as the average of the maximum and minimum possible efficient payoffs. In other words, kindness is based on the current choice and what could have occurred if different choices were taken. This allows the intentions of others to matter.

With perfect information, cooperation in pure strategies by second movers is only possible when the reversal probability for the first mover's choice is low (see proposition 4 in the Appendix A for details). In other words, the second mover will only cooperate when the first mover has high degree of control over her actions. When control is high and with sufficient concern about reciprocity, the second mover will cooperate if the first mover cooperates and defect if the first mover defects. Importantly, according to this model, the second mover will ignore the results of nature and condition their chose entirely on the first movers decision.

Under imperfect information, conditional cooperation is possible even when the reversal probability is high. As a result, second mover conditional cooperation may increase when the first mover's choice is uncertain (see proposition 4 in the Appendix A for details). Caution must be taken with this result because it relies on the sequential reciprocity equilibrium holding in the imperfect information setting. Under this equilibrium concept, second movers know with probability one the choice of the first mover. This is a strong assumption that may not hold.

Both inequity aversion (Fehr and Schmidt, 1999) and reciprocity (Dufwenberg and Kirchsteiger, 2004) can give quite different predictions in the sequential prisoner's dilemma with nature. It is possible that people may care about both outcomes and intentions. In Appendix A, a mixed concerns model is developed that combines concerns for inequity aversion (Fehr and Schmidt, 1999) and reciprocity (Dufwenberg and Kirchsteiger, 2004) into a single framework. The model captures predictions of both models but suggests an additional equilibrium under perfect information where the second mover only cooperates if the both the first mover and nature cooperate.

Importantly, the mixed concerns model and other combined models of outcomes and intentions like

Falk and Fischbacher (2006), predict that cooperation should increase as the reversal probability decreases. Reciprocity in these models depend on the control a person has over her choices. When first movers have greater control their decision to cooperate is viewed as a kinder action compared to when they have little control.

1.3.1 Individual Heterogeneity

Empathy

In the outcome-based and intention-based models of fairness it is assumed that each individual person can differ on how much they care about the different fairness concerns. While the models allow for individual heterogeneity it remains unclear why individuals differ in their concerns for fairness. One potential motivation for fair behavior may stem from individual capacity to empathize with others. In the *Theory of Moral Sentiments*, Adam Smith highlighted "compassion" or "fellow feeling" as the main factor in moral behaviors (Smith, 1790). This factor is now known as empathy and is essential in order for humans to understand others (Batson, 2011).

Empathic concern is the feeling of compassion or concern an individual has for the welfare of another person and the desire to help (Singer and Steinbeis, 2009; Batson, 2011). In the empathy-altruism hypothesis, empathic concern is proposed as the motivation for altruistic behavior. In order for empathic concern to be activated, an individual must perceive that another person is in need or value the welfare of that person (Batson et al., 2007). Once activated, empathic concern creates a desire to help that person. Empirical support suggests that empathic concern is an important component in altruistic behavior. Using survey evidence, empathic concern has been correlated with preferences for charitable giving (Bekkers, 2006; Ridinger, 2011), helping intentions (Kruger, 2003), and distributive justice (Ridinger, 2011). Additionally, empathic concern has been shown to be important in explaining behavior in dictator games (Bartels et al., 2013; Leliveld et al., 2012), and public good games (Batson et al., 1995; Oceja and Jimenez, 2007).

A few studies have suggested that empathic concern may be important in understanding cooperation in prisoner's dilemma games. In Batson and Moran (1999), female subjects played a one-shot prisoner's dilemma game with one-way communication. In the communication treatment, female subjects thought they were receiving written communication from the other player but instead the experimenters sent a note describing a negative personal experience. To induce different levels of empathy for the other player, subjects were asked to read the note objectively (low empathy) or try to imagine the situation from the other person's point of view (high empathy). Although the sample size was small, cooperation was higher when subjects read the note and even higher when adopting the viewpoint of the other person. Rumble et al. (2010) induced empathy in a similar way as Batson and Moran (1999) in a repeated prisoner's dilemma. Subjects believed they were playing with another player, but they were actually playing with a pre-programmed computer. Treatments varied whether the computer played tit for tat (no noise), tit for tat with noise (noise), and a noncooperative strategy. The results showed that the high empathy condition sustained high cooperation in the noisy condition, but not in the noncooperative strategy. Batson and Ahmad (2001) repeated the experiment in Batson and Moran (1999) except that female subjects were told that the other player had defected. The results showed that only 5% of subjects chose to cooperate in the no empathy condition, but 45% cooperated in the high empathy condition. These studies suggest that higher empathic concern may increase cooperation, but they cannot explain whether empathic concern drives increased cooperation through positive reciprocity, distributional concerns, or a combination of both.

In order for empathic concern to be activated, the target must be perceived as in need of help. In the sequential prisoner's dilemma with nature, cooperation by the first mover requires trust that the second mover will cooperate as well. By cooperating, the first mover makes themselves more vulnerable to exploitation. People who have higher empathic concern could be more likely to cooperate if the first mover cooperates. If the first mover defects, then it is more likely that the second mover will receive a lower payoff. First mover defection could be viewed negatively by second movers, and subsequently not activate empathic concern. If empathic concern is not activated, then there should be no significant difference in cooperation rates based on empathic concern. This suggests that higher empathic concern should correlate with positive reciprocity, but not with negative reciprocity.

Perspective Taking

The equilibrium concept assumed in the model of Dufwenberg and Kirchsteiger (2004) is quite strong as it requires individuals to have correct higher order equilibrium beliefs. Empirical studies of individual beliefs suggest that this assumption may be too strong for some individuals. When subjects choose in both roles, Blanco et al. (2014) found a "consensus effect" in a sequential prisoner's dilemma game where individual beliefs about the choices of others were biased towards one's own decision. In Dhaene and Bouckaert (2010), both first and second order beliefs were similar in the sequential prisoner's dilemma but were biased in the ultimatum game. Individual differences in the ability to predict others behavior is one potential explanation for these results.

Predicting others behavior in strategic environments appears to depend on perspective taking (Sher et al., 2014). Perspective taking is the ability to imagine or understand what another person is thinking or feeling (Batson, 2011). When an individual engages in perspective taking they may "put themselves in another person's shoes." Evidence suggests that perspective taking develops as children age and is deficient in individuals who have autism (Singer, 2009).

In the sequential prisoner's dilemma with nature, first movers must use perspective taking to attempt to predict what potential second movers will do. Second movers may use perspective taking to try to understand the meaning behind the observed actions of the first mover. For second movers who have low perspective taking, assuming that they hold correct first and second order beliefs may be too strong. Low perspective taking may hold beliefs that are more likely to be biased. In the mixed concerns model, the equilibrium beliefs determine the kindness of the first mover's action. If a subject has low perspective taking then they may be more likely to misinterpret the meaning of others' actions. In the perfect information case, the first mover can signal their intended action via their choice of cooperation or defection. If control is low, cooperation by the first mover is a less costly signal. As a result, even subjects who do not have "good" intentions may choose to cooperate hoping that second movers will think that they have "good" intentions and subsequently reward them. Individuals who have high perspective taking should be more likely to recognize that selfish first movers may be more likely to cooperate when control is low. As a result, they should be less likely to reciprocally cooperate when control is low compared to individuals who have low perspective taking. High perspective taking second movers may feel less guilty about defecting if the first mover cooperates because if control is low, then it is more likely they are defecting on a selfish cooperator. Similarly, when first mover control increases higher perspective taking should result in higher cooperation as they may be more likely to recognize the kindness of the first mover. When individuals have low perspective taking, observing higher cooperation by first movers may lead these second movers to think that people are being kind and as a result make these second movers more likely to cooperate when control is low.

1.4 Experiment

1.4.1 Experimental Design

A total of 246 students at a large public university participated in experimental sessions conducted in a computer laboratory. Students were recruited from a large subject pool. Recruitment to the subject pool took place through both classroom advertisements as well as through university emails. Prior to each experimental session, a random draw of students from the subject pool were sent an email about the upcoming session. Students then registered through the subject pool website. When registered students arrived at the experiment, they were randomly assigned to a computer terminal. No subject participated in more than one session.

The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). At the start of the experiment, subjects were randomly assigned to the role of first or second mover. This role stayed the same throughout the experiment and each round subjects were randomly matched with a different subject.[4] After reading instructions, subjects completed a quiz to test comprehension of the instructions[5]. Once finished, subjects played 20 rounds of a sequential prisoner's dilemma with nature. The role of nature was played by the computer. Each round the first mover could choose to cooperate or defect. After that choice, there was a random chance the computer reversed that choice. After both the first mover's choice and the results of the computer, the second mover then chose whether to cooperate or defect. To avoid potential framing effects subjects could choose A "cooperate" or B "defect."

Each subject participated in one of four possible treatments. Two sessions were conducted for each treatment for a total of eight sessions. The experiment used a within and between subjects design.

The reversal probability varied within subjects. Each treatment had a High Control condition and a Low Control condition. In the High Control condition, the chance the computer reversed the first movers choice was 10%.[6] For the Low Control condition, the reversal probability was 40%.Subjects received the High Control condition in the first ten rounds and the Low Control condition in the last ten rounds. Sessions were run in reverse to control for order effects. Each subject was told that the first ten rounds contained either the high control or low control. In order to allow between subject analysis, subjects were not told what the control would be in the last ten rounds until the start of the 11th round. The advantage of varying the control within subjects is that I can account for individual responses to the change in the reversal probability. Since it is assumed that utility functions vary by each individual i, using a purely between subject design creates the concern in smaller samples that differences in subject responses could be due to random differences in the utility functions between subjects and not necessarily due to the treatment variables.

[4]Due to not having exactly 40 subjects in each session there was some contamination in matching.
[5]A copy of the instructions given to subjects can be found in Appendix A.
[6]The reversal probability of 10% corresponds to figure 1.1 and 1.2 where $\theta = 0.1$.

The information varied between subjects. Each subject participated in either the Known or Uncertain treatment. In the Known treatment, subjects were told what the first mover chose and what the computer chose. In the Uncertain treatment, subjects were only told what the computer chose.

Subjects received a $7 show up payment and were paid based on three randomly selected rounds. The experiment lasted an average of 40 minutes each session. Subjects could earn anywhere from $10 to $19. The average amount earned by subjects was approximately $14. Table 1.1 gives information about the demographic characteristics of subjects by treatment. The average age, number of economic classes, and number of statistics classes are quite similar across treatments. Overall, 61% of subjects were female.

Table 1.1: Treatment Information

	High Control First		Low Control First		
	Known	Uncertain	Known	Uncertain	Total
Average:					
Age	20.34	20.53	20.56	20.24	20.42
Number of Economics Classes	1.18	1.33	1.26	1.53	1.32
Number of Statistics Classes	1.05	1.08	1.16	0.93	1.06
Take Home Earnings	13.66	13.83	13.76	13.84	13.77
Female (Fraction)	0.63	0.64	0.52	0.67	0.61
Number of Subjects	62	64	62	58	246

1.4.2 Psychometric Measures

After the experiment finished subjects completed a questionnaire containing demographic questions as well as psychometric tests designed to elicit levels of empathic concern, and perspective taking. To measure empathic concern and perspective taking, I used a subset of the interpersonal reactivity index (IRI) (Davis, 1983). The measure of empathic concern consisted of seven statements, for each statement subjects rated on a 5-point scale how well each statement described them. Examples of the statements are “I often have tender, concerned feelings for people less fortunate than me” and “When I see someone being treated unfairly, I sometimes don’t feel very much pity.” Similarly, the measure of perspective taking consisted of seven statements that subjects rated on a 5-point scale how well each statement described them. Examples of the statements are “I sometimes find it difficult to see things from the “other person’s” point of view” and “When I’m upset at someone, I usually try to “put myself in that person’s shoes”. Both sets of statements for empathic concern and perspective taking had strong internal consistency with Cronbach’s alphas of 0.73 and 0.70, respectively. Using subject responses, two variables representing empathic concern and perspective taking were derived using factor analysis. The eigenvalue for the empathic concern factor was 2.005 while the eigenvalue for the perspective taking factor was 2.009.

1.4.3 Hypotheses

Hypothesis 1: *(a) If subjects only care about inequity aversion, then there should be no differences in second mover cooperation as control or information changes. (b) If subjects care about reciprocity or mixed concerns, then cooperation rates should be larger in the Uncertain treatment relative to the Known treatment.*

Hypothesis 1 (a) directly from propositions 1 and 3 (see Appendix A for details). Changes in control and information do not change the outcomes at the end nodes of the game. As a result outcome-based models like inequity aversion predict no change in cooperation rates across the treatments. Hypothesis 1 (b) comes from the propositions 2, 4, and 5 (see Appendix A for details). If individuals care about reciprocity, the distribution of these concerns are similar across the treatments, and first mover cooperation is similar as information changes, then overall second mover cooperation rates should be higher in the Uncertain treatment compared to the Known treatment.

Hypothesis 2: *In the Known treatment:* (a) *Concern for reciprocity predicts that second mover co-*

operation will only occur if the first mover cooperates. (b) Inequity aversion predicts that second mover cooperation should only occur if the computer cooperates. (c) If subjects have mixed concerns, then in addition to the reciprocity and inequity aversion predictions, the model also predicts an additional equilibrium in the High Control condition where subjects only cooperate if the first mover and nature cooperates.

Hypothesis 2 (a) follows from propositions 2 and 4 while hypothesis 2 (b) follows from propositions 1 and 3. Hypothesis 2 (c) comes from proposition 6 (see Appendix A for details), and only occurs in pure strategies in the High Control condition.

Hypothesis 3: *If the first mover cooperates, then both reciprocity and the mixed concerns model predict that second mover conditional cooperation will be larger in the High Control condition compared to the Low Control condition.*

This hypothesis results from proposition 2, 4, 5, 6 and 7 (see Appendix A for details). In the reciprocity model, first mover cooperation should be viewed as kinder by second movers in the High Control condition compared to the Low Control condition. If first movers cooperate, then the reciprocity and mixed concerns model predict that second mover cooperation should be higher in the High Control condition.

Hypothesis 4: *In the Known treatment: (a) Second movers who have higher empathic concern should be more likely to cooperate if the first mover cooperated, but differences in empathic concern should have no effect on cooperation given the first mover defected. (b) Second movers with higher perspective taking should be less likely to conditionally cooperate in the Low Control condition compared to the High Control condition.*

1.4.4 Potential Econometric Issues

Due to subjects making repeated choices in the experiment, the use of standard ordinary least squares regressions is problematic as it is unlikely that the independence assumption will be met. In addition, the decision to cooperate or defect is a binary variable suggesting that non-linear regression is more appropriate. To deal with these issues, this paper uses both random effects probit regressions. The random effects probit model controls for random individual heterogeneity among subjects assuming there is no correlation between the individual error term and the independent variables. Although robust to other standard error assumptions, all regressions report clustered robust standard errors at the subject level. Where applicable, additional robustness checks were conducted using fixed effects logit (see Appendix A). The fixed effect logit controls for subject specific effects that do not change over time.

Another issue is that both empathy and perspective taking where measured after subjects participated in the experiment. Experimental conditions may have influenced how subjects answered the empathy and perspective taking. If this is true, then any correlation between the treatments and empathy and perspective taking could be due to the post elicitation of the measures. Table S1 and the subsequent analysis using nonparametric tests in Appendix A shows that scores on empathic concern and perspective taking where not significantly different from each other by treatment and condition. This suggests that the treatments themselves do not seem to have led subjects to answer differently to the empathy and perspective taking measures.

One potential concern is that due to possibly high correlation between empathy and perspective taking that including both these terms in the same regression may create multicollinearity. Additional regressions including only empathy or only perspective taking, show that the coefficient estimates are largely similar. While there is high correlation between the two variables, the estimates for their effects on cooperation appear to be unaffected by their simultaneous inclusion in the regression.

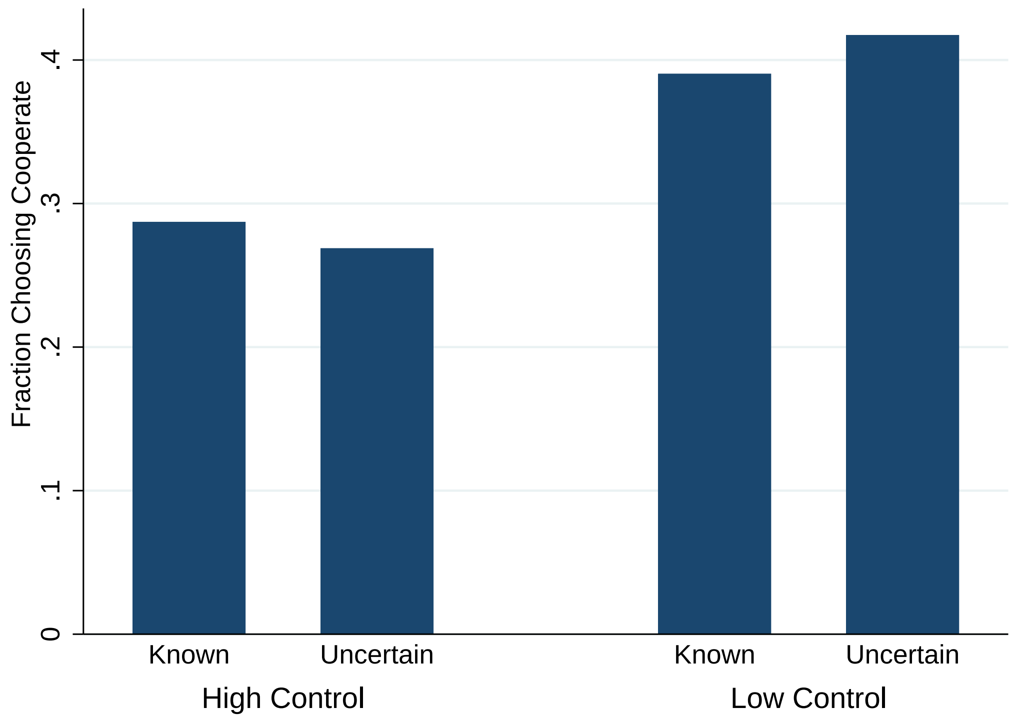


Figure 1.3: Average First Mover Cooperation by Treatment (First 10 Rounds)

1.5 Results

Result 1: *Second mover cooperation was higher in the Low Control and Uncertain condition.*

Figure 1.3 shows the average first mover cooperation by treatment restricted to the first 10 rounds.[7] First movers cooperated more often in the Low Control condition compared to the High Control Condition. Cooperation between the Known and Uncertain treatments appears to be similar. Figure 1.4 shows the average second mover cooperation by the computer choice. Second movers cooperated less often when the first movers choice was known compared to uncertain. This suggests that knowledge of the first movers choice mattered contrary to what is predicted by outcome-based fairness models. Looking at the Known Treatment, figure 1.5 shows the average second mover cooperation rates for the different paths of play. Clearly, cooperation was higher when the computer cooperated

[7]This restriction is done for clarity of presentation as the results are similar when using the all 20 rounds.

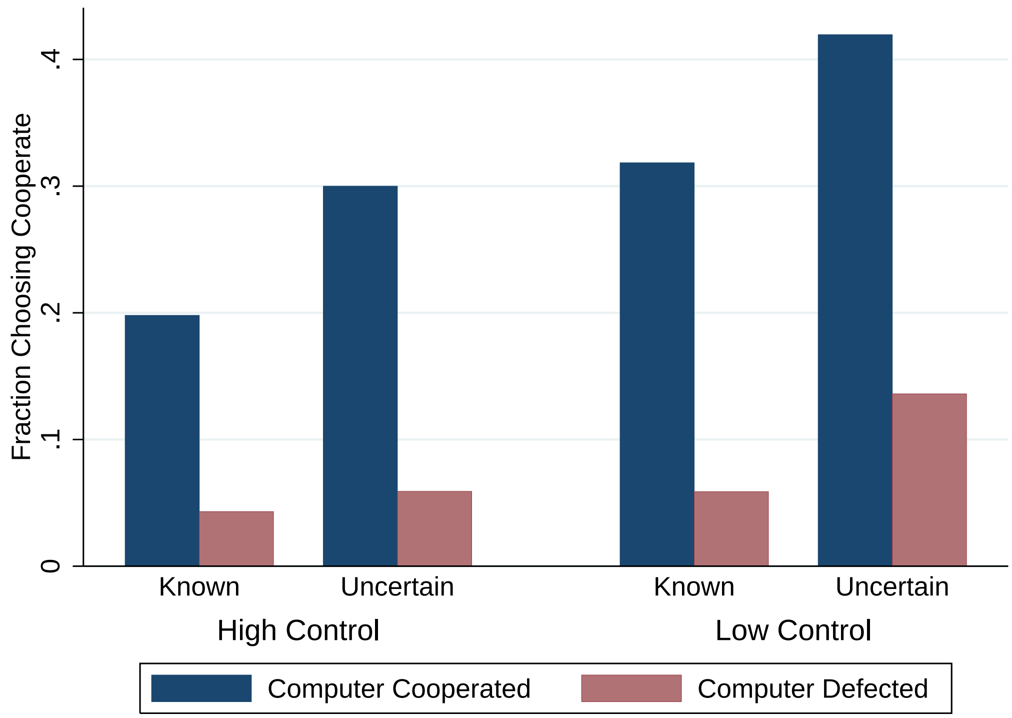


Figure 1.4: Average Second Mover Cooperation by Treatment (First 10 Rounds)

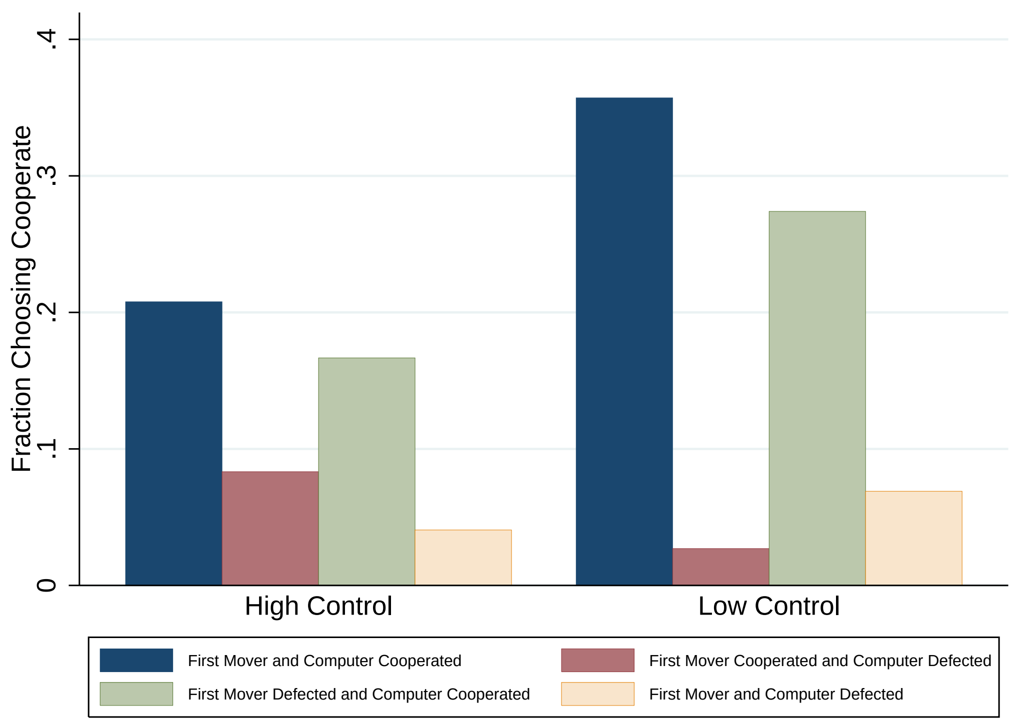


Figure 1.5: Average Second Mover Cooperation in Known Treatment (First 10 Rounds)

Table 1.2 gives results from random-effects probit regressions using the data from the first 10 rounds. The restriction allows a between subjects analysis, and the results suggest that cooperation by the first mover was significantly higher in the Low Control condition. Higher cooperation by first movers when control is low could be due to the fact that cooperation is less costly since there is a high chance that their choice will be reversed. In the uncertain treatment, first movers could avoid having subjects learn about their choices, but they were still aware of how their choice influenced the second movers. [8] Interestingly, first mover behavior was not influenced by whether their intended choice would be known or unknown to the second mover. Second movers cooperated more often in the Uncertain treatment and Low Control condition. These effects seem to be additive since the interaction was not significant[9]. Despite the lack of change in cooperation by first movers in the Uncertain treatment, second movers were more likely to cooperate in the Uncertain treatment. Similar to Cox and Deck (2006), it appears that subjects gave first movers the benefit of the doubt.

Result 2: *Inequity aversion, reciprocity, and mixed concerns are unable to fully explain the experimental results.*

Purely reciprocal second movers in the Known treatment should cooperate only if the first mover cooperated and ignore the computers choice. Purely inequity averse players should cooperate at the same rates for all treatments and cooperate only if the computer cooperates. The results from table 1.3 show that neither models of pure reciprocity nor pure inequity aversion alone can explain the observed results in the Known treatment. Subjects were clearly drawn to the Pareto superior outcome. However, in the Known treatment, second movers were more likely to choose the Pareto superior outcome if the first mover cooperated. This result suggests that second movers were influenced by the intentions of the first mover.

Whenever the first mover's choice is reversed there are potentially competing norms for fair minded

[8]This is similar to the "plausible deniability" treatment in the dictator game experiment by Dana et al. (2007).

[9]In addition, these results are robust to including age, number of economic classes, number of statistics classes, gender, and political views.

Table 1.2: First Mover and Second Mover Cooperation (First 10 Rounds)

	First Mover		Second Mover	
	Cooperation	Predicted Probability	Cooperation	Predicited Probability
Uncertain	−0.01	−0.00	0.47*	0.06*
	(0.13)	(0.04)	(0.21)	(0.03)
Low Control	0.42**	0.15**	0.46*	0.06*
	(0.13)	(0.04)	(0.21)	(0.03)
First Mover and Computer cooperated			1.30***	0.26***
			(0.15)	(0.04)
First Mover cooperated and Computer defected			0.10	0.01
			(0.24)	(0.03)
First Mover defected and Computer cooperated			1.20***	0.27***
			(0.17)	(0.05)
Female	0.10	0.04	−0.21	−0.03
	(0.13)	(0.05)	(0.21)	(0.03)
Intercept	−0.11		−1.52***	
	(0.18)		(0.27)	
N	1230	1230	1230	1230
ρ	0.23***	0.23***	0.45***	0.45***
Model χ^2	70.76	70.76	135.90	135.90

Predicted probabilities represent a discrete change from 0 to 1. Cluster robust standard errors at the subject level in parentheses. Results are from random-effects probit models that includes round fixed effects. Dependent variable is equal to 1 if player cooperated and equal to 0 otherwise.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

subjects. One hypothesis is that when faced with conflicting norms people will be more likely to select the norm that coincides with their own self-interest (Bicchieri, 2006). If first movers cooperated and the computer defected, cooperation by second movers was insignificant. While reciprocity suggests subjects should cooperate, inequity aversion predicts that subjects will defect. While this fits the hypothesis that people will select the norm that makes them personally better off, this does not seem to be the case when the first mover defected and computer cooperated. Here cooperation by second movers was significant. Reciprocity predicts that subjects should defect while inequity aversion suggests people will cooperate. While there was less cooperation

Table 1.3: Second Mover Cooperation by Treatment

	Known		Uncertain	
	High Control	Low Control	High Control	Low Control
First Mover and Computer cooperated	1.61***	1.84***	1.42***	1.50***
	(0.25)	(0.28)	(0.23)	(0.29)
First Mover cooperated and Computer defected	−0.03	−0.32	0.07	0.33
	(0.63)	(0.46)	(0.64)	(0.33)
First Mover defected and Computer Cooperated	1.02**	1.42***	1.75***	1.57***
	(0.31)	(0.29)	(0.33)	(0.27)
Low Control First	0.29	1.44*	−0.16	2.11**
	(0.50)	(0.61)	(0.53)	(0.69)
Female	−0.15	0.18	−0.63+	−0.21
	(0.34)	(0.40)	(0.36)	(0.51)
Intercept	−1.74***	−3.01***	−1.19**	−2.65***
	(0.44)	(0.62)	(0.44)	(0.70)
N	589	620	578	578
ρ	0.51	0.62	0.50	0.68
Model χ^2	59.61	66.98	61.45	65.73
Hypothesis Tests				
Inequity Aversion (Prob> $\chi^2(1)$)	0.07+	0.07+	0.31	0.78
Reciprocity (Prob> $\chi^2(1)$)	0.01**	0.00***	0.04*	0.00***

Cluster robust standard errors at the subject level in parentheses. Hypothesis for Inequity Aversion is that cooperation given computer cooperated is the same regardless of first mover's choice. Hypothesis for Reciprocity is that cooperation given first mover cooperated is the same regardless of computer's choice. Results are from random-effects probit regressions with round fixed effects.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

if the first mover defected, cooperation was much higher than we would suspect if people choose between competing norms by selecting the norm that maximizes their own payoff.

In the uncertain case, we fail to reject the predictions of inequity aversion but the reciprocity predictions are rejected. This test of the reciprocity model is more of a test of the SRE concept, because the concept requires that the second mover knows for certain that the first mover cooperated even though they only observed the computer defecting. This suggests that when players have imperfect information the SRE concept could be too strong.

From table 1.3 it appears that order mattered in the experiment. The variable Low Control First is equal to one if subjects received the Low Control condition in the first ten rounds. In both the Known and Uncertain treatments, subjects who started the experiment with Low Control cooperated at higher rates compared to subjects who received the Low Control second. This suggests that there was some path dependence in overall cooperation rates depending on which condition subjects received first. Despite this, the direction of change is consistent with the result that subjects cooperated more often when control was low.

Result 3: *In the Known treatment, if the first mover cooperated and the computer cooperated, second mover cooperation was higher in the Low Control condition compared to the High Control condition.*

Clearly, second movers were influenced by the different treatments and by the path of play. The mixed concerns model allows players to care about both what the first mover chose and the results from the computer. In the Known treatment, the mixed-concerns model predicts that when control is high and the first mover cooperates, then cooperation should be higher compared to the low control treatment. This prediction is not supported by the results. In round 1 of the Known treatment, given that both the first mover and the computer cooperated, second mover cooperation in the high control treatment was 17.6% compared to 52.9% in the low control treatment. These cooperation rates are significantly different from each other (Wilcoxon rank-sum test, z=-2.121, p=0.034). This result is puzzling because theoretical predictions from reciprocity suggest that people should interpret cooperation by the first mover when control is high as kinder than cooperation when control is low. However, it appears that subjects responded in the opposite way as the model predicts. Since all end node payoffs for both players were kept constant, it appears that the difference is primarily through how individuals were influenced by the reversal probability.

Result 4: *(a) In the Known and Low Control Treatment, second movers with lower perspective taking were more likely to cooperate. (b) In the Known Treatment, second movers with higher empathic concern cooperated more often if the first mover cooperated.*

Perspective taking could be important in how others interpret or try to understand the actions of others. If individuals care about others' intentions, then perspective taking may be an influential factor in determining the intentions of others. Table 1.4 shows the results looking at the role of perspective taking on cooperation. In columns (1) and (2) , differences in perspective taking do not appear to have influenced overall cooperation across all treatments. When the regressions were restricted to the Known Treatment, perspective taking by itself is not significant. However, when perspective taking is interacted with the Low Control condition the interaction is significant. Higher perspective taking was associated with lower cooperation rates in the Low Control treatment.

Figure 1.6 classifies individuals with lower than median perspective taking as low perspective and higher than median perspectve taking as high perspective. In the high control treatment average conditional cooperation by high perspective takers is higher than low perspective takers. However, in the low control condition this is reversed. The regression results from table 1.5 show that the interaction term is significant both when the first mover cooperated, and when the first mover defected. Figure 1.7 plots the predicted probabilities from table 1.5 by scores in perspective taking for subjects in the High and Low Control conditions. When the first mover cooperated there is a clear decline in cooperation as scores in perspective taking increase. This decline actually crosses the high control treatment suggesting that individuals who have a high degree of perspective taking may have been more likely to cooperate in the High Control condition. Figure 1.8 plots the mean difference in predicted probability between the High and Low Control conditions by perspective taking. This graph shows that individuals who score low on perspective taking were significantly more likely to cooperate in the Low Control condition compared to low perspective takers in the High Control condition. While figure 1.4 suggested that high perspective takers may have been more likely to cooperate in the High Control condition this difference is not significant.

It is possible that subjects mistook first mover cooperation for kindness in the Low Control condition. One potential explanation for subjects with low perspective taking is that they just did not

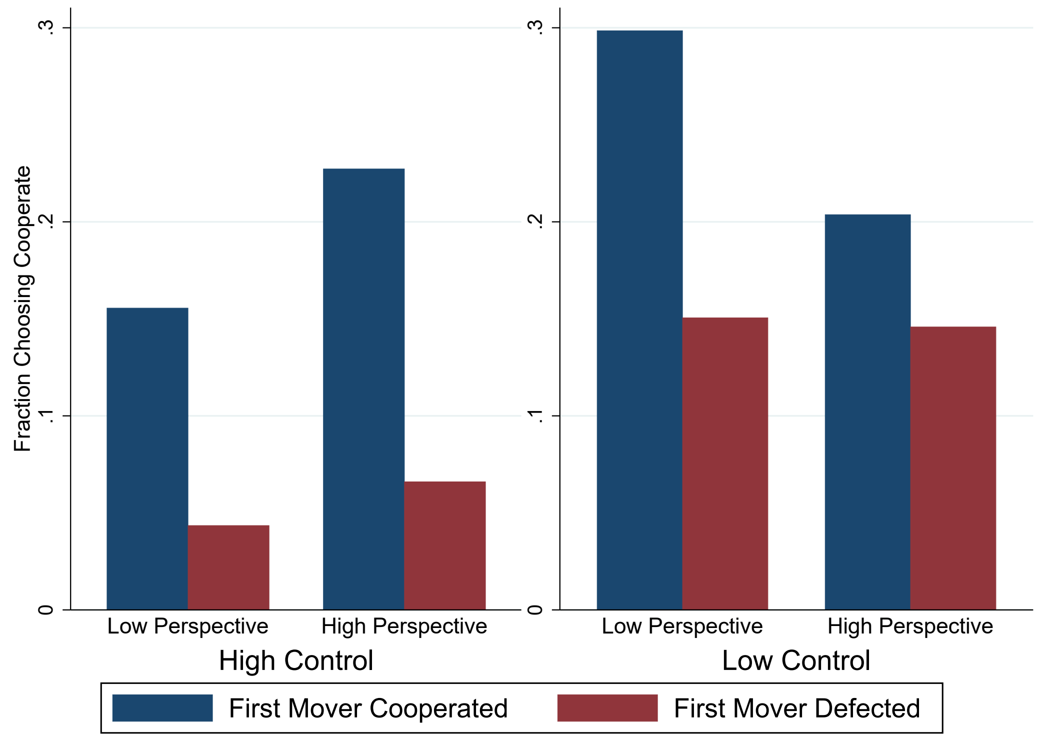


Figure 1.6: Average Second Mover Conditional Cooperation in Known Treatment by Perspective Taking (First 10 Rounds)

understand what they were doing. While this is possible, it seems unlikely because there was no effect from perspective taking in the High Control condition. One explanation for this could be that intentions were much easier to understand in this situation. When the first mover cooperated, most of the time the first movers choice matched the computer result. This may have made it easier for subjects to understand the meaning of the first movers choice despite differences in perspective taking abilities. Recent research has suggested that people have an automatic intuition to cooperate (Rand et al., 2012). Low perspective takers may have found difficulty in inferring the meaning behind first mover cooperation in the Low Control condition. As a result, they may have been more likely to go with their gut instinct to cooperate.

Empathic concern could potentially lead to increased cooperation through it's influence on altruism. From table 1.5, columns (1) and (2) support this view as subjects with higher empathic concern were more likely to cooperate overall. Columns (3) and (4) in table 1.5 show that higher

Table 1.4: Second Mover Cooperation with Empathy and Perspective Taking (First 10 Rounds)

	All Treatments		Known	
	(1)	(2)	(3)	(4)
	Second Mover Cooperation	Second Mover Cooperation	Second Mover Cooperation	Second Mover Cooperation
Empathic Concern	0.30^{*}	0.31^{*}	0.41^{+}	0.39^{+}
	(0.13)	(0.13)	(0.23)	(0.22)
Perspective Taking	0.04	0.18	−0.21	0.14
	(0.13)	(0.15)	(0.23)	(0.28)
Low Control	0.55^{*}	0.56^{*}	0.45	0.38
	(0.21)	(0.22)	(0.34)	(0.33)
Low Control X Perspective Taking		−0.37		-0.87^{**}
		(0.24)		(0.34)
Uncertain	0.52^{*}	0.52^{*}		
	(0.21)	(0.21)		
First Mover and Computer cooperated	1.33^{***}	1.33^{***}	1.43^{***}	1.42^{***}
	(0.18)	(0.18)	(0.28)	(0.28)
First Mover cooperated and Computer defected	0.04	0.03	−0.38	−0.39
	(0.24)	(0.24)	(0.48)	(0.50)
First Mover defected and Computer Cooperated	1.23^{***}	1.23^{***}	1.14^{***}	1.17^{***}
	(0.19)	(0.19)	(0.28)	(0.28)
Female	−0.34	−0.32	−0.21	−0.07
	(0.21)	(0.22)	(0.34)	(0.31)
Intercept	-1.53^{***}	-1.57^{***}	-1.54^{***}	-1.63^{***}
	(0.29)	(0.29)	(0.47)	(0.46)
N	1200	1200	610	610
ρ	0.43	0.42	0.50	0.46
Model χ^2	137.51	138.86	64.17	65.65

Cluster robust standard errors at the subject level in parentheses. Regressions are restricted to the first 10 rounds. Results are from random-effects probit regressions with round fixed effects.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

empathic concern is associated with increase cooperation in the known treatment. The coefficient for empathic concern is not significant when the regressions are restricted to the uncertain treatment. This suggests that higher empathic concern may only increase cooperation if subjects can view the actions of the first mover.

Table 1.5: Conditional Cooperation in Known Treatment with Empathy and Perspective Taking

	First Mover Cooperated Second Mover Cooperation	First Mover Defected Second Mover Cooperation
Empathic Concern	0.64^{*}	0.25
	(0.31)	(0.21)
Perspective Taking	0.07	0.09
	(0.41)	(0.23)
Low Control	0.37	0.34
	(0.53)	(0.30)
Low Control X	-0.97^{*}	-0.68^{*}
Perspective Taking	(0.48)	(0.34)
Computer cooperated	2.18^{*}	1.14^{***}
	(0.89)	(0.27)
Female	-0.16	-0.10
	(0.47)	(0.36)
Intercept	-2.77^{*}	-0.90^{+}
	(1.13)	(0.50)
N	209	306
ρ	0.61	0.34
Model χ^2	21.95	29.68

Cluster robust standard errors at the subject level are in parentheses. Regressions are restricted to the first 10 rounds. Results are from random-effects probit regressions with round fixed effects.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.6 examines the influence of empathic concern in the Known Treatment condition on the first movers choice. If first movers cooperated, individuals with higher empathic concern were more likely to cooperate. If first movers defected, no significant difference in cooperation occurred based on differences in empathic concern. This supports the hypothesis that first mover cooperation activates empathic concern leading to cooperation. Defection by first movers does not activate empathic concern, and as a result empathic concern does not influence behavior. It appears that empathic concern may be an important factor in individual desire to reward others for "good" intentions.

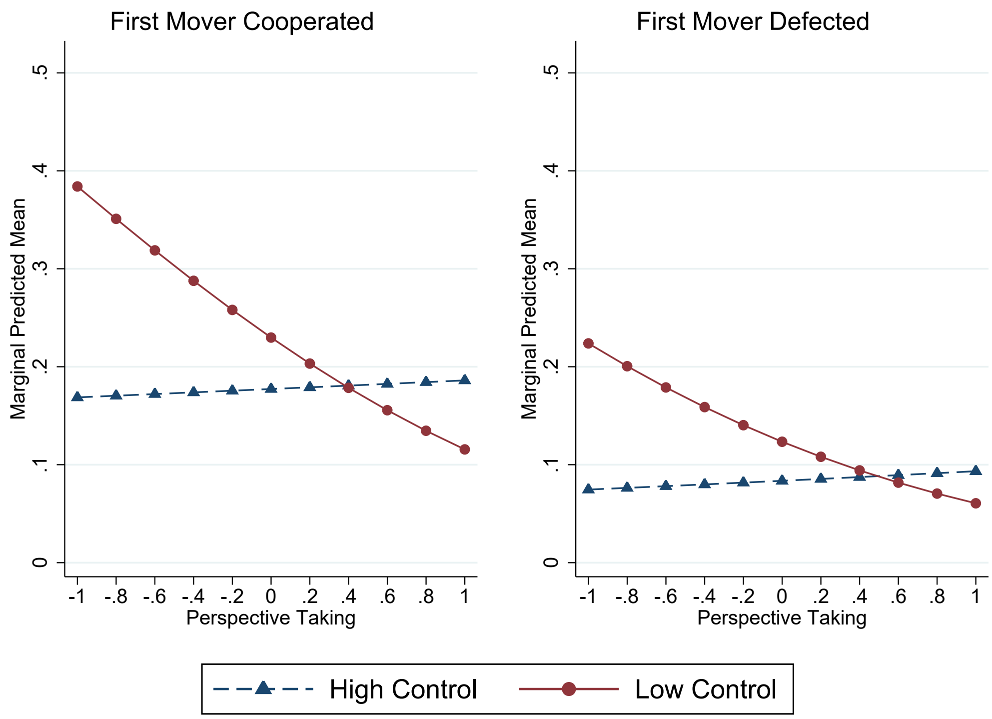


Figure 1.7: Predicted Probability by Perspective Taking in Known Treatment (First 10 Rounds)

1.6 Conclusion

This paper has shown that cooperation rates in the sequential prisoner's dilemma with nature are greater than predicted by pure self-interest. The failure of pure-self interest to explain behavior in a wide range of one-shot experimental games has led many researchers to suggest fairness concerns as a potential explanation for the empirical results (Camerer, 2003). Theoretical models of fairness can be classified into two types: outcome-based and intention-based. Outcome-based models of fairness assume that people care about fair distributions (Bolton and Okenfels, 2000; Fehr and Schmidt, 1999). These models assume that intentions are not relevant for predicting behavior. The results from this experiment suggest that outcomes matter, but purely consequentialist models cannot fully explain subject behavior. Intention-based models capture the idea that people are reciprocal, preferring to be kind to people who are kind to them and punish people who are unkind to them (Dufwenberg and Kirchsteiger, 2004; Rabin, 1993). In this experiment intentions mattered

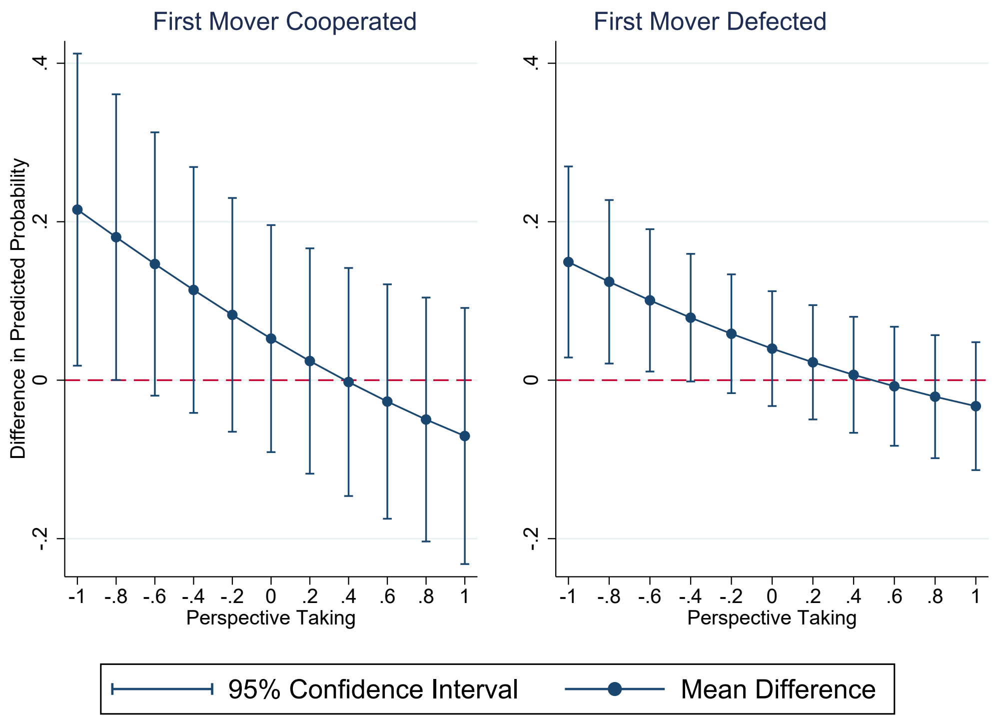


Figure 1.8: Mean Difference in Predicted Probability between Low Control and High Control conditions by Perspective Taking (First 10 Rounds and Known Treatment).

as well, but current models of intentions failed to explain the concerns for fair outcomes.

The mixed motives model combined both inequity aversion and reciprocity into a single framework. While able to account for concerns about both outcomes and intentions, the model was unable to explain the increased second mover cooperation in the Low Control treatment. The results highlight that how individuals perceive others' intentions is still an open question. It is not possible for people to know with certainty the true intentions of another person. Despite this, people potentially use the observed actions of others to infer intentions. These intentions could influence how people respond to others' behavior and appear to be important in understanding cooperation.

One important aspect for understanding the attribution of intentions may be perspective taking. Differences in the ability to take the viewpoint of others was shown to be important in explaining increased cooperation by second movers when control was low. While concerns for fairness are important, both empathic concern and perspective taking could be significant factors in explaining subject behavior in games. Second movers who scored higher on empathic concern where more likely to cooperate. When the first movers choice was known, higher empathic concern increased cooperation when the first mover cooperated but not when the second mover defected. This demonstrates that empathic concern is different from negative reciprocity, since subjects with higher empathic concern were not more likely to defect if the first mover defected. Instead, empathic concern is thought to motivate altruism. This altruistic motivation is activated when the first mover cooperates leading to increased cooperation. Future research should include measures of empathic concern and perspective taking to investigate potential relationship in other games and environments.

Chapter 2

Ownership, Punishment, and intentions in real effort bargaining

2.1 Introduction

Decades of experimental research on the ultimatum game has consistently found that behavior often deviates from what is predicted by individual payoff maximization (Camerer, 2003; Guth and Kocher, 2014). Insights learned from these studies have expanded knowledge of how people bargain and led to numerous game-theoretic models that can explain individual behavior.[1] For example, Falk et al. (2003) showed using a ultimatum game with restricted offers that rejection rates of the same offer differed depending on what the proposer could have chosen. This result suggests that individuals may be motivated not only by the distribution of payoffs but also the intentions of the proposer. When subjects are given the money to bargain by the experimenter proposers often offer even splits, and half the time responders reject offers lower than 20% (Camerer, 2003).

[1]These models of fairness focus on individual concerns about the distribution of payoffs (Bolton and Okenfels, 2000; Fehr and Schmidt, 1999) or intention-based reciprocity (Dufwenberg and Kirchsteiger, 2004; Rabin, 1993) or a combination of both (Charness and Rabin, 2002; Falk and Fischbacher, 2006; Ridinger, 2015a).

However, individual behavior can differ when subjects produce the surplus from real effort or investment (Ganter et al., 2001; Gachter and Riedl, 2005; Garcia-Gallego et al., 2008; Fischbacher et al., 2009; Bediou et al., 2012; Feng et al., 2013; Karagozoglu and Riedl, 2014) or if they hold entitlements over the surplus (Hoffman et al., 1994; Leliveld et al., 2008). One explanation for these results is that individuals are influenced by social norms (Bicchieri, 2006). These norms or rules about behavior can vary depending on the context of the situation. The production of the surplus or feelings of entitlement may change the relevant norm that individuals use in their decisions.

It is important to note that in the vast majority of ultimatum game experiments, individuals have bargained over surpluses given to them by experimenters. Often in real world bargaining, individuals negotiate over a surplus that was previously produced. According to standard theory, whether or not the negotiating parties produced the surplus themselves should have no impact on individual behavior. This is due to prior costs of production being sunk at the start of negotiations. However, recent research has shown that production before negotiations can strongly influence individual behavior in bargaining (Ganter et al., 2001; Gachter and Riedl, 2005; Garcia-Gallego et al., 2008; Fischbacher et al., 2009; Bediou et al., 2012; Feng et al., 2013; Franco-Watkins et al., 2013; Karagozoglu and Riedl, 2014) . Research on joint production (i.e. when both parties in the negotiations take part in the production of the surplus) has found that individual behavior is more in line with equity concerns (Ganter et al., 2001; Gachter and Riedl, 2005; Fischbacher et al., 2009; Bediou et al., 2012; Karagozoglu, 2012; Feng et al., 2013; Franco-Watkins et al., 2013; Karagozoglu and Riedl, 2014). Joint production occurs in many real-world environments including wage negotiations, government budgets, and team production. The joint production process can potentially generate important norms like entitlements (i.e. beliefs about what a person should receive) or equity that may impact bargaining behavior.

One particular set of norms that may impact bargaining behavior has to do with ownership. Norms about ownership are learned early in childhood development. These rules of ownership are so important that conflict over ownership is thought be one of the most common occurrences in young children's lives (Hay and Ross, 1982). Children between the ages of 2 and 4 often ascribe ownership to the first possessor of an object (Friedman and Neary, 2008). In play sessions, 20 and 30 month old children who claimed prior ownership over objects were more likely to instigate and win property conflicts with children who did not have prior ownership (Ross, 2013). Research with adult populations in ultimatum and dictator games has suggested that norms of ownership may have strong effects on individual behavior (Hoffman et al., 1994; Leliveld et al., 2008). Despite the importance of ownership norms still little is known how prior ownership influences individual behavior when both parties feel entitled to a share of the surplus due to joint production.

In this paper, I ask: How does prior ownership influence individual behavior in bargaining over a jointly produced surplus? Do changes in the effectiveness of responder punishment affect adherence and enforcement of ownership norms? This paper uses a laboratory experiment to investigate how prior ownership, punishment, and intentions influence individual bargaining behavior over a jointly produced surplus. In the experiment, each subject is paired with another subject in a real effort task, where their combined efforts jointly produce a surplus. This surplus is allocated to the subject that preforms better on the task. In the next stage, the two subjects bargain over the surplus in a series of four mini-ultimatum games with one game randomly selected for payment. In a 2x2 between subjects design, treatments varied both the property rights and the strength of punishment by the responder. In the Responder Surplus treatment, the surplus was allocated to responders. In the Proposer Surplus treatment, the surplus was assigned to proposers. In the Equal Threat treatment, rejection by the responder left both subjects with \$0. In the Unequal Threat treatment, rejection by the responder gave the proposer \$2 and the responder \$0.

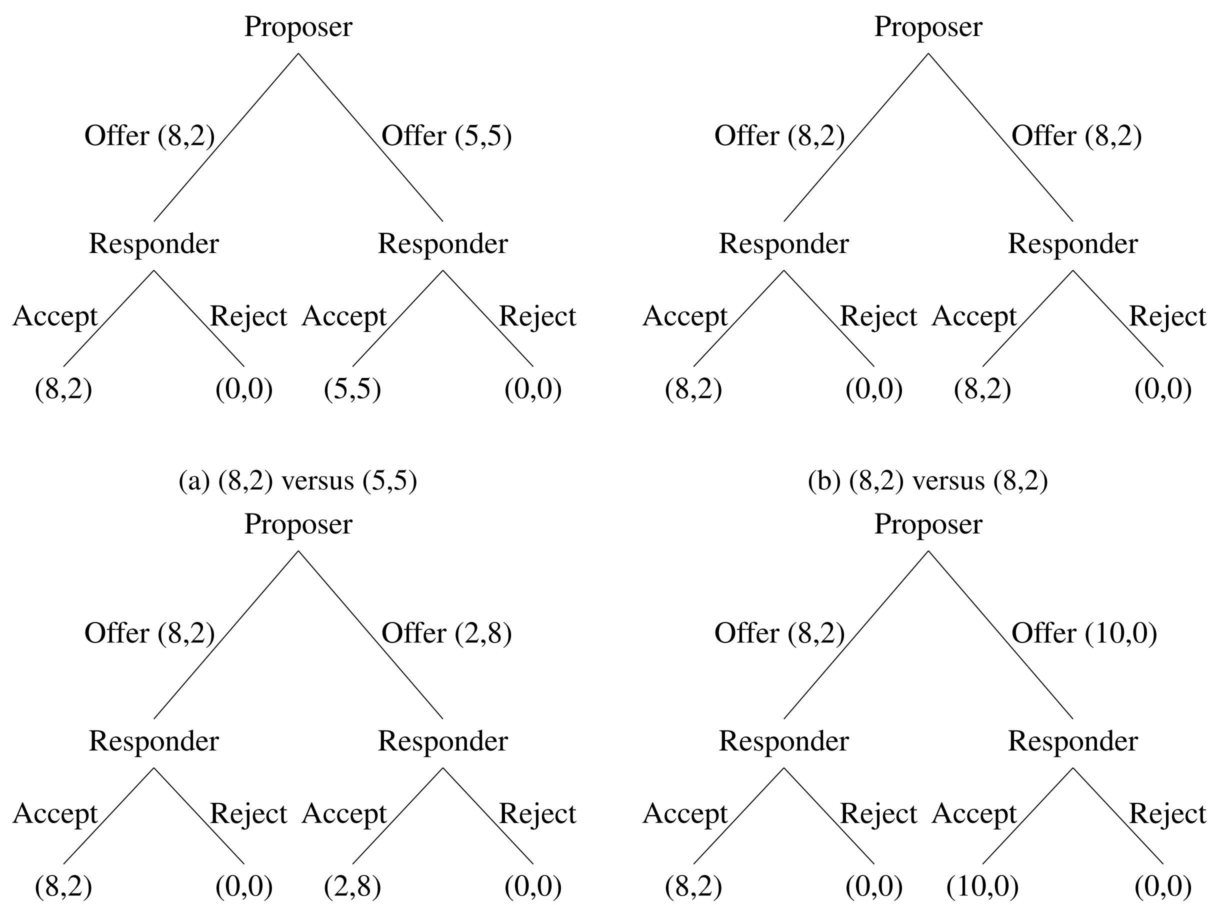
I find that proposers respect prior responder ownership when the responder has a strong ability to punish, but not when punishment is weak. Responders respect prior proposer ownership when their ability to punish is weak, but reject at high rates when they have strong punishment. Even when controlling for differences in individual efforts, the respect of ownership by responders in the weak punishment condition still remains. An independent measure of rule-following shows that individuals who are more likely to follow the rules are more likely to respect prior ownership. In addition, in varying ownership and the strength of punishment, the distribution of subjects that conditionally accept unequal offers changes. This adds an additional explanation for the empirical finding that intentions seem to matter in some cases but not in others(Charness, 2004; Charness and Levine, 2007; Falk et al., 2008, 2003; Falk and Kosfeld, 2006; McCabe et al., 2000, 2003; **?**).

This paper contributes to previous research on joint production with real effort by examining the importance of prior ownership and punishment in behavior in bargaining. In addition, by using mini-ultimatum games the results shed light on how the alternatives available influence both proposer and responder behavior. This paper also adds new insight to the understanding of social norms in bargaining by showing that an independent measure of rule-following is predictive of bargaining behavior.

Previous experiments examining the importance of ownership or entitlements in ultimatum games include Hoffman et al. (1994) and Leliveld et al. (2008). In Hoffman et al. (1994), subjects answered quiz questions for the right to be the proposer. Subjects who earned the right, made less equal offers to responders. Despite this, rejection rates by the responders were relatively unchanged compared to treatments where the right to be the proposer came from random draw. This experiment in this paper differs from Hoffman et al. (1994) in that subjects will be completing a real-effort task that is designed to be less reliant on intelligence or cognitive ability. In this experiment, subjects participate in a real-effort task that involves moving sliders across a screen. Differences in intelligence or backgrounds are unlikely to be correlated with performance. In addition, this experiment will allow responders to have prior ownership as well, to see if offers by the proposer and rejection rates are influenced . In Leliveld et al. (2008), subjects did not earn the endowment but treatments varied whether the endowment was common property, given to the responder, or given to the proposer. Offers by proposer changed significantly depending on the initial assignment of the endowment. Deception was used as all subjects played the role of the proposer which could have biased the results. While proposers seem to be influenced by prior ownership, it is not clear what the effect is on responders' behavior.

Using survey evidence, Franco-Watkins et al. (2013) present subjects with different scenarios of ultimatum games. Some subjects received a standard ultimatum game and others received ultimatum games where effort was involved in the production of the surplus. The authors found that effort changed both the offers and subjects' judgment of the fairness of offers. Bediou et al. (2012) examined the importance of equality or equity by having subjects answer math problems to jointly produce a surplus prior to an ultimatum game. They found that when proposers produced less than half of the surplus they tended to offer equal splits, but when they produced more than half they made offers that where between equal splits and the equitable share. This paper builds on this literature by examining how prior ownership effects behavior when both parties produced the surplus.

Falk et al. (2003) studied the importance of distributions and intentions by conducting an experiment in which subjects played four mini-ultimatum games illustrated in figure 2.1. In all the games, proposers could select an (8,2) offer which gave \$8 for herself and \$2 for the responder. The possible alternative offers available to the proposer varied across the games. While the results showed some support for distributional fairness concerns, the results also showed that rejection rates of the (8,2) offer changed systematically depending on the alternatives. Responders were much more likely to reject when the proposer could have offered an even split (5,5) compared to when the alternative was (2,8). This result suggests that people care about more than just the final distribution of payoffs. This paper adds to our understanding of the results in Falk et al. (2003) by studying behavior in the mini-ultimatum game when the surplus is jointly produced by real effort.



(a) (8,2) versus (5,5)

(b) (8,2) versus (8,2)

(c) (8,2) versus (2,8)

(d) (8,2) versus (10,0)

Figure 2.1: Mini-Ultimatum Games with Strong Punishment

2.2 Experimental Design

2.2.1 Treatments

A total of 198 students from a large public university participated in an experiment at a computer laboratory. Subjects were recruited from a large and diverse subject pool. Prior to each experimental session, a randomly selected group of students from the subject pool were sent an email about participating in the upcoming session. Students then registered through the subject pool website. At the day of the experiment, registered students were randomly assigned to a computer terminal. No subject participated in more than one experimental session. To facilitate data collection, the experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

Each subject was paired randomly with another subject. Each subject then had to complete a real effort task. The real effort task used was the slider task introduced by Gill and Prowse (2012). In the slider task, each subject must move sliders across the screen. The task is designed such that it is difficult for a subject to finish the task before the time runs out. The advantage of the slider task is that it is not cognitively difficult which reduces the potential for differences in performance confounding potential treatment effects. Another advantage of the task is that it is not exciting. This makes the task costly for subjects to complete. This is important because the goal of using the task was for subjects who receive the endowment to feel like they earned it due to their own costly efforts.

In pairs, the surplus was created via joint production between a subject and his or her paired partner. Subjects were not told the production function, but were aware that the money produced was based on their combined efforts. To facilitate data analysis, the surplus produced was either $10 or $0 based on an effort threshold. If the combined number of number of sliders was greater than one, then $10 were produced. Otherwise, $0 were produced. In the pair, the subject who completed more sliders received the surplus. In the case of tie, one subject was chosen at random with a 50% probability. After completing the task, subjects were told how many sliders they completed, how many sliders their partner completed, the surplus they received, and the surplus their partner received.

After completing the real effort task, each subject remained paired with their previous partner. Subjects then had to bargain over the surplus created in a series of mini-ultimatum games (Falk et al., 2003). In the mini-ultimatum games the potential offers by the proposer are restricted to two options as seen in figure 2.1. In each game, the proposer can always offer to keep $8 for herself and $2 for the responder. Each game differs in the alternative offer available to the responder. Each player was either the proposer or the responder and this role remained the same for all four games. The results from one of the four games was randomly selected for payment which better mimics one-shot games. The strategy method was used in order to collect data on conditional strategies for each individual responder. To reduce the potential for bias, subjects were not told about the results of the four games until the end of the experiment. The order of the mini-ultimatum games was randomized at the subject level to help control for potential order effects.

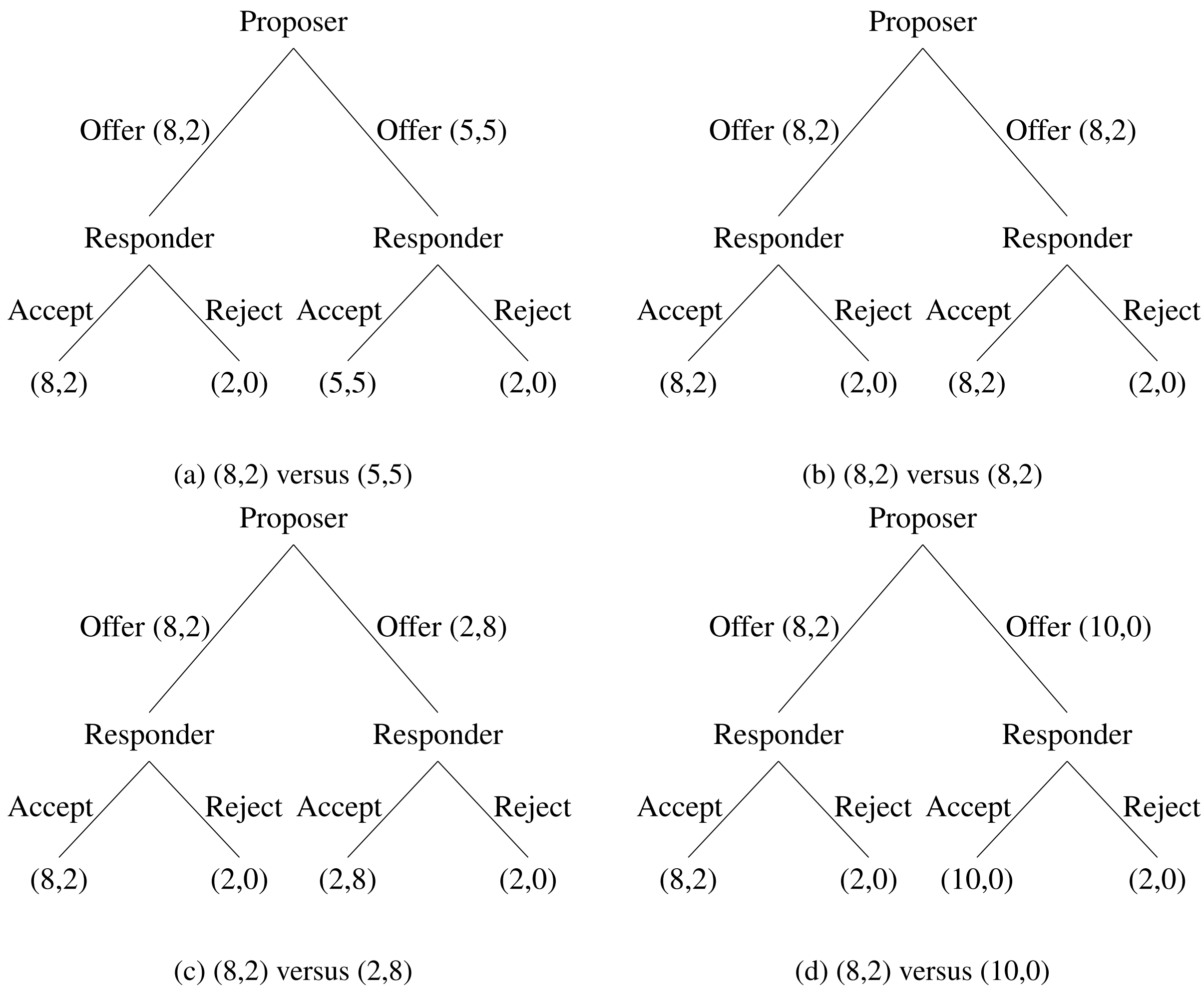
Treatments followed a 2x2 factorial design by varying the strength of punishment by the responder and whether the proposer or responder received the surplus. Each subject participated in exactly one of four possible treatments. In the Strong Punishment treatment, rejection by the responder left both players with a payoff of $0 as seen in figure 2.1. In the Weak Punishment treatment, rejection by the responder gave the proposer $2 and the responder $0 as seen in figure 2.2. In the Proposer Surplus treatment, the proposers received the surplus from the real effort task. In the Responder Surplus treatment, the responders received the surplus from the real effort task.

At the end of the experiment, each subject learned the results of her interaction and completed a questionnaire asking demographic questions.[2]

2.2.2 Rule-Following Task

In all treatments, each subject participated in the rule-following task designed to measure individual sensitivity to follow a costly rule or norm (Kimbrough and Vostroknutov, 2015). In the task, each subject controlled a stick figure that walked from the left side to the right side of the screen. As the figures walked, they automatically stopped at a traffic light. There were a total of five traffic lights. At any time, subjects could start walking again by pressing the walk button. After five seconds the light would turn green. Subjects began with $7 and were told that for each second they spent on the task they would loose $0.07. Walking without waiting at any of the lights took a total of 24 seconds. Subjects who did not wait would loose at least $1.68. Subjects who waited the full

[2]Although not discussed in this paper subjects also completed the following tasks: Reading the Mind in the Eyes test (Baron-Cohen et al., 2001), Interpersonal Reactivity Index(Davis, 1983), Guilt and Shame Proneness Scale(Cohen et al., 2011), Mach IV Scale (Christie et al., 1970), and the Cognitive Reflection Test (Frederick, 2005).



(a) (8,2) versus (5,5)

(b) (8,2) versus (8,2)

(c) (8,2) versus (2,8)

(d) (8,2) versus (10,0)

Figure 2.2: Mini-Ultimatum Games with Weak Punishment

five seconds at each light lost $3.08. In the instructions, subjects were told "The rule is to wait at each stop light until it turns green." An example of what a typical subject would see in the task can be seen in figure 2.3.

The rule-following task simply created a situation where a person could follow a rule that is costly to themselves. The time that subjects took to complete the task can be viewed as a measure of an individuals sensitivity to following rules or social norms. Kimbrough and Vostroknutov (2015) show that this measure is good predictor of subject behavior in wide range of experimental games. In addition, they show that the rule that subjects are told appears to be what is leading some subjects to wait. Without the rule, the majority of subjects ignore the rule, suggesting that the difference in

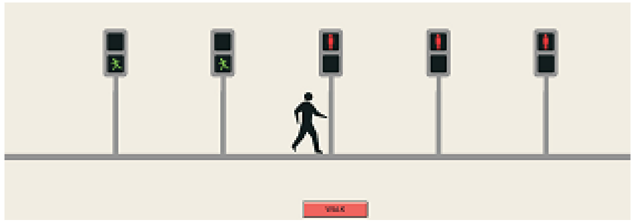


Figure 2.3: Example of Rule-Following Task

Table 2.1: Summary Statistics by Treatment

	Strong Punishment		Weak Punishment		
	Proposer Surplus	Responder Surplus	Proposer Surplus	Responder Surplus	All Treatments
Average					
Age	20.70	20.10	20.48	20.19	20.37
Female	0.69	0.66	0.65	0.65	0.66
Number of Economics Classes	1.48	1.06	2.50	1.21	1.55
Number of Statistics Classes	1.15	1.04	1.89	0.92	1.24
Native English Speaker	0.59	0.59	0.66	0.62	0.61
Overall Subject Earnings	14.11	14.40	15.03	14.41	14.47
Number of Subjects	54	50	46	48	198

behavior was driven by the stated rule. Kimbrough and Vostroknutov (2015) note that one potential criticism of this measure is that it is simply a measure of the experimental demand effect. If this is the case, then subjects who wait longer in the task are doing so because the experimenter told them to do it. This demand effect is precisely what we are interested in measuring, because we are interested in individual differences in the propensity to follow a given costly rule or norm.

2.2.3 Descriptive Statistics

Table 2.1 presents the summary statistics by each treatment. Two sessions were run for each treatment with each session lasting approximately one hour. The average take home earnings for subjects across all treatments was \$14.47. The majority of subjects in the treatments were female with similar fractions across the different treatments. Table 2.2 gives the average of the different effort variables by treatment. Both Proposer and Responder Efforts are calculated as the number

Table 2.2: Summary Statistics of Effort Variables by Treatment

	Strong Punishment		Weak Punishment	
	Proposer Surplus	Responder Surplus	Proposer Surplus	Responder Surplus
Average				
Proposer Effort	43.00	26.88	40.83	30.08
Responder Effort	17.19	43.60	26.87	40.46
Proposer Effort Ratio	0.74	0.35	0.62	0.42
Responder Effort Ratio	0.26	0.65	0.38	0.58

of sliders that subjects completed in the slider task[3].

2.3 Predictions

In this section, a simple model based on Bicchieri (2006) and Kessler and Leider (2012) is introduced to help motivate hypotheses on how norms may influence individual behavior in bargaining.[4] Suppose that two individuals are negotiating over a positive surplus X. Let the share of the surplus received by player i be denoted as x_i and the share of player j be denoted as x_j where $x_i + x_j = X$. Define the norm based utility of an individual as:

$U_i(x_i, x_j, \theta_i) = x_i - \gamma_i \cdot g(\theta_i(h) \cdot X - x_i)$ if $\theta_i(h) \cdot X > x_i$,

$U_i(x_i, x_j, \theta_i) = x_i$ otherwise.

The term $\theta_i(h) \in [0,1]$ represents what player i's believes the norm or norms suggest about the share of the surplus they should receive at information set h.[5] This gives the norm-based share of the surplus for player i as $\theta_i(h) \cdot X$. If player i receives an actual share x_i that is less than the

[3] The Proposer Effort Ratio is equal to Proposer Effort divided by the sum of the Proposer and Responder Effort. The Responder Effort Ratio is calculated similarly.

[4] Both Bicchieri (2006) and Kessler and Leider (2012) present models of norm based utility, but differ in a key way. Bicchieri (2006) presents a linear model where individuals care about norms when they receive a payoff that is less than payoff they would receive based on the norm. Kessler and Leider (2012) presents a more general model where individuals prefer to choose an action that is closer to a commonly known normative action.

[5] This assumption is important in extensive form games as the norm may differ depending on path of play.

norm-based share, then player i's utility is decreased. The monotonically increasing function g captures the effect on utility due to norm violation. The term $\gamma_i \geq 0$ captures the degree in which player k cares about payoffs that deviate from what she expected to receive based on the norm.

Hypothesis 1: *Proposers should be less likely to offer (8,2) when the responder earned the right and when there is strong punishment.*

Hypothesis 2: *Responders should be more likely to reject the (8,2) offer when the proposer earned the right and when there is strong punishment.*

Norms, captured by $\theta_i(h)$, can potentially influence what people think they should receive. These norms activated by social cues may change what player i thinks should be the share she receives. Receiving the entire surplus prior to the bargaining stage may have activated norms of ownership. This effect could occur in individuals that received the surplus by increasing the share they think they should receive. For individuals that did not receive the share prior to bargaining, the norms of ownership may lead to expectations of a smaller share. If norms of ownership matter, then equilibrium predictions from the model generate hypotheses 1 and 2 (see Appendix B for details).

Hypothesis 3: *Responder rejection rates of (8,2) offers should change depending on the alternative available to the proposer.*

Changes in the alternatives available to the proposer can influence how individual responders view identical offers (Falk et al., 2003). One interpretation of these results is that different alternatives may signal a different norm (Bicchieri, 2006). When the proposer can choose an equal split of (5,5) but instead chooses an unequal split of (8,2), then the responder may view that as a violation of the norm. However, when the choices are (8,2) or (10,0), and the proposer offers (8,2), then the responder may view the proposer as honoring the norm. Since the share $\theta_i(h)$ is a function of the information set h, the model can account for changes in the norm depending on the path of play. If alternatives available change the norm, then rejections should differ depending on the alternative available.

Hypothesis 4: *Higher relative efforts by responders (proposers) should lead to increased rejection (offer) of the (8,2) division.*

In the experiment, subjects jointly produce the surplus by their own efforts. Hence, subjects may feel that the equitable share they receive should be based on both their own and their partner's efforts. The accountability principle captures this norm, predicting that people may judge an outcome to be fair if the allocation is proportional to relevant variables that each party can control (Konow, 2000). For example, each person can control the effort that she puts in the task. The accountability principle suggests that what constitutes a fair offer in the ultimatum game should be proportional to efforts by each person. If the accountability principle is an important norm in this experiment, then individual efforts should be predictors of offers and rejections in the ultimatum game.

Hypothesis 5: *(a) Proposers who were more likely to follow the rule-following task should be less likely to offer (8,2) when the responder received the right (b) Responders who followed the rule in the rule following task should be less likely to reject (8,2) offers when the proposer earned the right.*

The rule-following task attempts to measure individual sensitivity to follow norms. In the above model this heterogeneity can occur if subjects differ on their individual sensitivity to care about norms or if they differ on the norm. If we assume that the norms are similar across subjects, then the behavior in the rule-following task can give us an estimate of the individual sensitivity to norms. While many individuals may hold similar norms, it is unlikely that all subjects will hold the same norms. As a result, using the rule-following task may be an imprecise estimate of the individual sensitivity to norms in the above model. Despite this, the rule-following task does give a measure of the overall tendency of individuals to sacrifice their own payoff in order to follow a rule or norm. Individuals who sacrifice more are likely people who care more about following the rules or norms. This suggests that proposers who follow the rules in the rule following task may be more likely to select a responder favorable offer when the responder received the right. Similarly, responders who follow the rules may be less likely to reject proposer favorable offers when the proposer earned the right.

2.4 Results

Result 1: *If the responder earned the right and had strong punishment, then proposers were less likely to offer (8,2) when there was a more responder favorable alternative.*

Figure 2.4 plots the fraction of subjects offering (8,2) across the treatments. In the Strong Punishment treatment, overall proposers were more likely to offer $8 for themselves and $2 for the Responder (i.e. (8,2)) when they earned the right (Wilcoxon rank sum, z=2.604, p=0.009). Proposers offered $2 for themselves and $8 for the responder more often when the Responder earned the right (Wilcoxon rank sum, z=2.108, p=0.0351). In the weak punishment treatment, proposer offers of (8,2) did not seem to be influenced by who received the right (Wilcoxon rank sum, z=-0.094, p=0.925). In addition, no statistical difference for prior ownership was found between offers of (8,2) when the alternative was (2,8).

This result is consistent with the interpretation that proposers were more likely to respect ownership of the responder when the responder had a stronger ability to punish. Proposers were more likely to offer (2,8) to the responder when the ability to punish was strong compared to weak (Wilcoxon rank sum, z=-2.670, p=0.008). One explanation for this result could be do to changes in proposer expectations about the likelihood of rejection by the responder. Proposers were more willing to select the unequal offer in favor of the responder when they thought it was more likely that the responder would punish them. It is also likely that proposers did not choose the (2,8) offer when punishment is weak, because in that case they are guaranteed to get $2 regardless of the responders choice. As a result, it is profit maximizing to choose (8,2) if they believe that there is some small chance that a responder will accept.

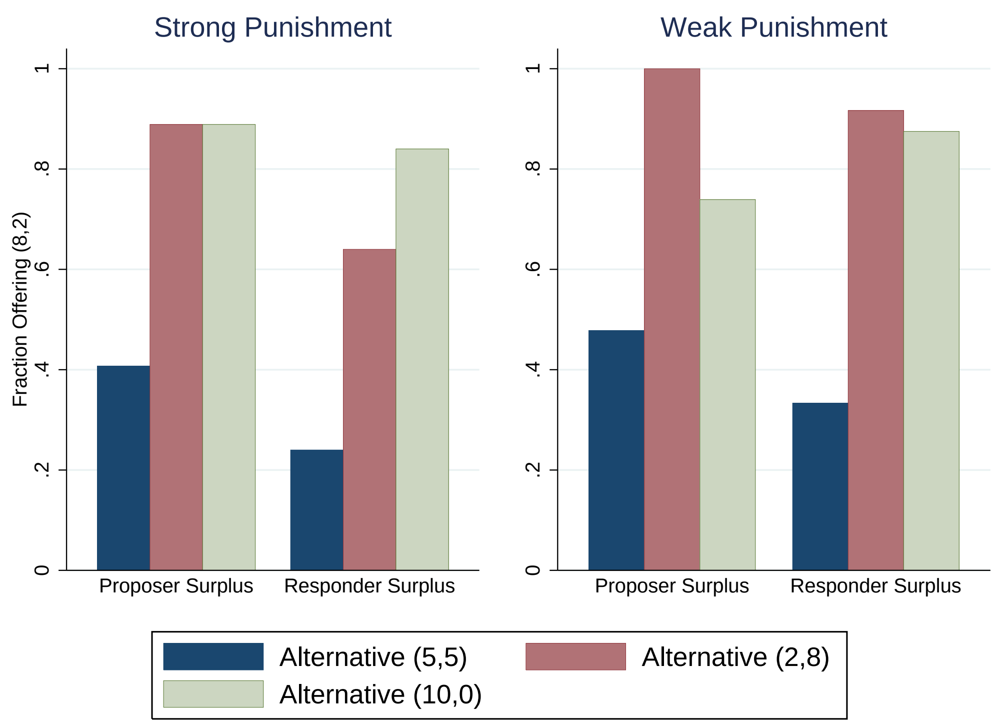


Figure 2.4: Average Offer of (8,2) by Treatment

Result 2: *(a) In the Strong Punishment treatment, rejection rates of (8,2) by responders were similar between the Proposer and Responder Surplus treatments. (b) In the Weak Punishment treatment, responders were less likely to reject if the proposer earned the right compared to if the responder earned the right.*

Figure 2.5 gives the fraction of responders rejecting the (8,2) offer by treatment. In the Strong Punishment treatment, rejection rates of the (8,2) offers were similar between the property right treatments (Wilcoxon rank sum, z=-0.094, p=0.925). In the Weak Punishment treatment, responders were more likely to reject the (8,2) if they earned the right compared to if the proposer earned the right (Wilcoxon rank sum, z=-1.945, p=0.051).

One interesting result, is that rejection rates by responders were quite high when punishment was strong. In Falk et al. (2003), the rejection rate of an (8,2) offer when the alternative was (10,0) was 9%. When the responder earned the right and had a strong ability to punish, the rejection

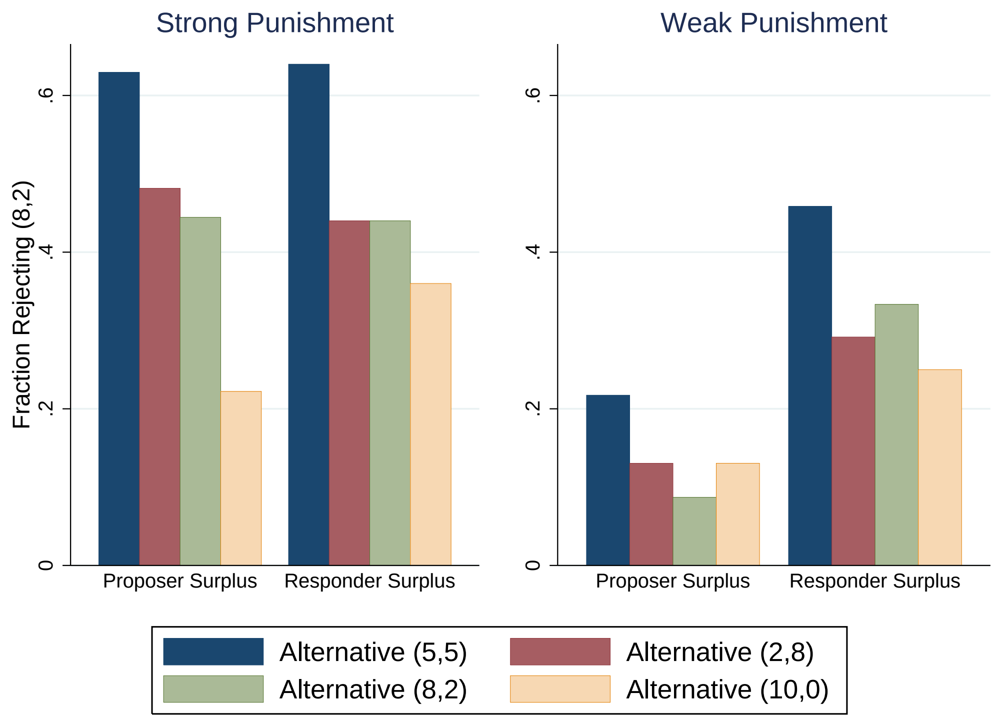


Figure 2.5: Average (8,2) rejections by Treatment

rate of (8,2) when the alternative was (10,0) was 36%. This is a big difference between these two experiments. Other than subject pool differences, it seems that either prior ownership, the real effort production of the surplus, or both led to this difference. This result highlights the fact that concerns for fairness are quite malleable and depend greatly on the context of the situation.

These results illustrate the difficulty with fairness models that focus on only outcomes(Bolton and Okenfels, 2000; Fehr and Schmidt, 1999), intentions(Dufwenberg and Kirchsteiger, 2004; Rabin, 1993), or both (Falk et al., 2003; Ridinger, 2015a). The fairness models capture certain fairness norms that seem to matter, but ignore salient factors that influence behavior. It seems that changes in prior ownership and punishment can change the appearance of different fairness concerns. This could be due to these external factors changing the relative weights people place on different fairness concerns. As a result, these concerns are not fixed and are highly context dependent.

Result 3: *(a) When there was strong punishment, there was a larger fraction of subjects who conditionally accepted or rejected the (8,2) offer based on the alternative when the proposer earned the right. (b) When punishment was weak, there was a larger fraction of subjects who conditionally accepted or rejected the (8,2) offer based on the alternative when the responder earned the right.*

In the Strong treatment, subjects were significantly more likely to conditionally accept/reject when the proposer had the right (Wilcoxon rank-sum test, z=1.692, p=0.091). In the Weak treatment, subjects were significantly more likely to conditionally accept/reject when the responder had the right (Wilcoxon rank-sum test, z=-1.724, p=0.085). Conditional acceptance or rejection of the (8,2) offer can be interpreted as a concern for intentions. Prior ownership changed the distribution of people who appear as if they are influenced by intentions. When the responder had the right and punishment was strong, a large fraction of subjects rejected offers of (8,2) no matter the alternative. In this case, subjects may have viewed this offer as unfair no matter what the proposer could have chosen.

The fraction of responders who conditionally accept increased when the proposer had the right and the responder had strong punishment. It appears that this different context led more subjects to take into account what the proposer could have chosen, instead of rejecting any unequal offer. When punishment was weak, conditional acceptance was still high when the responder had the right. The fact that a small amount of subjects are classified as concerned about the intentions of others in the weak punishment and proposer right case does not mean that most subjects did not care about the other players intentions. What this result suggests is that the observed behavior, that people respond to the same offer differently depending on what alternative is available, is sensitive to the context or norms of the situation.

Result 4: *(a) In the Weak Punishment treatment, even controlling for efforts, responders were less likely to reject unequal offers if the proposer earned the right.*

One potential issue is that while prior ownership seemed to matter it is unclear whether this was

Figure 2.6: Subject Choice Categories by Treatment

due to differences in individual efforts or who received the surplus prior to bargaining. The following analysis shows that the assignment of the right is still significant for responders even when controlling for individual efforts. This suggests that who had prior ownership may have influenced responder behavior beyond accounting for differences in effort. As a result, it lends support that idea that prior ownership may influence the propensity of responders to enforce the norm and change how responders view unequal offers.

Table 2.3 regressed rejection rates of the (8,2) offer on the treatments and efforts. Column (1) suggests that subjects were less likely to reject in the Weak Punishment treatment, but more likely to reject when the responder had prior ownership. Column (2) shows that the increase in rejections in the Responder Surplus treatment is driven primarily from subjects in the Weak Punishment treatment. Columns (3) and (4) accounts for the effort ratio between the responder and proposer. Higher values mean that the responder put in higher effort relative to the proposer. The coefficient

Table 2.3: Predicting Rejection of (8,2) Offer by Treatment and Effort

	(1) Reject (8,2)	(2) Reject (8,2)	(3) Reject (8,2)	(4) Reject (8,2)
Responder Surplus	0.26^{+}	0.03	0.19	−0.23
	(0.14)	(0.19)	(0.21)	(0.28)
Weak Punishment	-0.64^{***}	-0.93^{***}	-0.64^{***}	-1.02^{***}
	(0.14)	(0.21)	(0.14)	(0.22)
Responder Surplus X Weak Punishment		0.54^{+}		0.68^{*}
		(0.28)		(0.30)
Responder Effort Ratio			0.23	0.67
			(0.49)	(0.53)
Alternative (5,5)	0.73^{***}	0.74^{***}	0.73^{***}	0.74^{***}
	(0.19)	(0.19)	(0.19)	(0.19)
Alternative (2,8)	0.31	0.31	0.31	0.31
	(0.20)	(0.20)	(0.20)	(0.20)
Alternative (8,2)	0.28	0.28	0.28	0.28
	(0.20)	(0.20)	(0.20)	(0.20)
Female	-0.32^{*}	-0.30^{*}	-0.31^{*}	-0.27^{+}
	(0.14)	(0.14)	(0.15)	(0.15)
Age	0.06	0.05	0.06	0.04
	(0.05)	(0.05)	(0.05)	(0.05)
Native English Speaker	0.15	0.16	0.15	0.16
	(0.14)	(0.14)	(0.14)	(0.14)
Intercept	-1.63^{+}	−1.29	-1.68^{+}	−1.37
	(0.98)	(1.00)	(0.99)	(1.00)
N	396	396	396	396
Pseudo R^2	0.09	0.10	0.09	0.10
χ^2	48.37	52.07	48.59	53.64

Each column presents random effects probit regressions. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

for effort ratio is not significant. Column (4) shows that the results are similar to column (2) except the coefficients increase in magnitude and significance.

Table 2.4 gives the results of regressing offers of (8,2) by proposers on both treatment and effort variables. The regressions are restricted to include only the alternatives (5,5) and (2,8) in order to

Table 2.4: Predicting Offer of (8,2) by Treatment and Effort

	(1) Offer (8,2)	(2) Offer (8,2)	(3) Offer (8,2)	(4) Offer (8,2)
Responder Surplus	-0.65^{**}	-0.75^{**}	-0.55^{+}	-0.67
	(0.22)	(0.28)	(0.32)	(0.43)
Weak Punishment	0.53^{*}	0.41	0.54^{*}	0.44
	(0.21)	(0.31)	(0.21)	(0.33)
Responder Surplus X Weak Punishment		0.24		0.19
		(0.42)		(0.45)
Proposer Effort Ratio			0.33	0.21
			(0.75)	(0.80)
Alternative (5,5)	-1.56^{***}	-1.56^{***}	-1.56^{***}	-1.56^{***}
	(0.22)	(0.22)	(0.22)	(0.22)
Female	0.07	0.07	0.08	0.07
	(0.22)	(0.22)	(0.22)	(0.22)
Age	-0.09	-0.09	-0.10	-0.10
	(0.07)	(0.07)	(0.07)	(0.07)
Native English Speaker	-0.09	-0.09	-0.08	-0.08
	(0.22)	(0.22)	(0.22)	(0.22)
Intercept	3.18^{*}	3.22^{*}	2.99^{*}	3.10^{*}
	(1.41)	(1.41)	(1.47)	(1.49)
N	198	198	198	198
Pseudo R^2	0.27	0.27	0.27	0.27
χ^2	70.17	70.48	70.36	70.55

Each column presents random effects probit regressions. Standard errors in parentheses.
Regressions results are restricted to (8,2) offers when alternatives were (5,5) and (2,8).
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

focus on offers of (8,2) when the alternative was better for responders. In column (1), the coefficient for Responder Surplus is negative and significant suggesting that proposers were less likely to offer (8,2) when the responder had prior ownership. The coefficient for Weak Punishment is positive and significant in column (1), but is no longer significant in column (2) when the interaction term with Responder right is added. Columns (3) and (4)

One potential issue with the regression in tables 3 and 4 is that the Responder Surplus treatment and effort ratio are correlated. To deal with this, additional regressions (available upon request) were conducted examining the effect of the effort ratio on rejections and offers of (8,2) restricted to the Responder Surplus and Proposer Surplus treatment. These results suggest that the individual differences in efforts do not seem to have had a significant impact on rejections or offers of (8,2). This suggests that the reason for the differences in behavior across the treatments may be due to norms of ownership and not differences in effort.

Result 5: *(a) Responders who followed the rule in the rule-following task were less likely to reject the (8,2) offer when the proposer received the right (b) Proposers who followed the rule in the rule-following task were less likely to offer (8,2) when the responder received the right.*

Table 2.5 shows regression results looking at how behavior in the rule following task influenced rejections of the (8,2) offer. Column (1) shows that individuals who waited longer at the lights in the rule following task were significantly less likely to reject unequal offers when the proposer had prior ownership. Column (3) suggests that waiting time did not have a significant effect when the responder had prior ownership. Columns (2) and (4) look at subjects who never waited at any of the five stop signs. The results show that these subjects were significantly more likely to reject the unequal offers in both the Proposer and Responder right treatments. Together these results lend support that responders who care about following norms are more likely to respect proposer ownership.

Table 2.6 looks at how rule following behavior influences proposers decision to offer (8,2). Column (1) shows that waiting longer was not correlated with (8,2) offers when the proposer had prior ownership. In column (3), individuals who waited longer were less likely to offer the (8,2). Columns (2) and (4) show that while the coefficient is positive, no significant difference in behavior was found in proposers who never waited at any of the lights. While not definitive, these results suggest that individuals who were more likely to follow the costly rule were more likely to respect responder ownership.

Table 2.5: Predicting Rejection of (8,2) Offer by Rule Following

	(1) Proposer Surplus	(2) Proposer Surplus	(3) Responder Surplus	(4) Responder Surplus
Waiting Time	-0.03^{+}		-0.01	
	(0.02)		(0.01)	
Never Waited		1.13^{**}		1.07^{*}
		(0.41)		(0.42)
Weak Punishment	-1.05^{***}	-1.05^{***}	-0.32	
	(0.24)	(0.24)	(0.20)	
Alternative (5,5)	0.86^{**}	0.87^{**}	0.69^{*}	0.82^{**}
	(0.29)	(0.29)	(0.27)	(0.28)
Alternative (2,8)	0.49^{+}	0.50^{+}	0.18	0.48^{+}
	(0.29)	(0.30)	(0.27)	(0.29)
Alternative (8,2)	0.34	0.35	0.24	0.35
	(0.30)	(0.30)	(0.27)	(0.29)
Responder Effort Ratio	0.62	0.67	0.92	-0.74
	(0.77)	(0.77)	(0.79)	(0.65)
Female	-0.27	-0.26	-0.37^{+}	-0.44^{+}
	(0.24)	(0.24)	(0.20)	(0.23)
Age	-0.03	-0.06	0.06	0.01
	(0.07)	(0.07)	(0.08)	(0.07)
Native English Speaker	-0.29	-0.16	0.73^{**}	-0.22
	(0.20)	(0.21)	(0.23)	(0.20)
Intercept	0.96	0.63	-2.27	-0.58
	(1.72)	(1.47)	(1.63)	(1.41)
N	200	200	196	200
Pseudo R^2	0.16	0.18	0.10	0.10
χ^2	40.57	44.80	25.57	24.76

Each column presents random effects probit regressions. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Result 6: *(a) Responders who never waited at the light during the rule-following task were less likely to conditionally accept (8,2) offer (b) Proposers who waited longer in the rule-following task were more likely to conditionally offer (8,2).*

Table 2.6: Predicting Offer of (8,2) by Rule Following

	(1) Proposer Surplus	(2) Proposer Surplus	(3) Responder Surplus	(4) Responder Surplus
Waiting Time	−0.00		−0.05*	
	(0.02)		(0.02)	
Never Waited		0.03		0.62
		(0.54)		(0.54)
Weak Punishment	−0.14	−0.14	0.14	0.24
	(0.30)	(0.31)	(0.33)	(0.33)
Alternative (5,5)	−1.08***	−1.08***	−1.82***	−1.74***
	(0.27)	(0.27)	(0.32)	(0.31)
Proposer Effort Ratio	−0.13	−0.14	−1.80	−1.95
	(0.97)	(0.97)	(1.34)	(1.30)
Female	0.02	0.01	0.22	−0.02
	(0.28)	(0.28)	(0.42)	(0.41)
Age	−0.05	−0.05	−0.10	−0.07
	(0.08)	(0.08)	(0.12)	(0.12)
Native English Speaker	0.01	0.01	0.05	0.05
	(0.29)	(0.29)	(0.32)	(0.31)
Intercept	2.08	2.04	4.83+	3.09
	(1.83)	(1.80)	(2.65)	(2.42)
N	100	100	98	98
Pseudo R^2	0.13	0.13	0.33	0.30
χ^2	16.76	16.75	44.33	40.42

Each column presents random effects probit regressions. Standard errors in parentheses.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7 examines how conditional acceptance and rejection of the (8,2) offers were influenced by behavior in the rule following task. Column (1) shows that waiting longer in the task did not significantly influence whether a responder conditionally accepted the (8,2) offer. Column (2) shows that individuals who never waited were much less likely to conditionally accept or reject (8,2) offers compared to subjects who waited at one or more lights. This result is driven by the fact that subject who never waited were much more likely to reject all (8,2) offers no matter the proposer alternative. This result fits with the argument that changes in the alternatives may change

Table 2.7: Predicting Conditional Rejection of (8,2) Offer by Rule Following

	(1) Conditionally Reject (8,2) Offer	(2) Conditionally Reject (8,2) Offer
Waiting Time	0.03	
	(0.02)	
Never Waited		-1.33^{*}
		(0.63)
Responder Surplus	-0.85	-0.79
	(0.57)	(0.57)
Weak Punishment	-1.04^{*}	-1.05^{*}
	(0.42)	(0.42)
Responder Surplus X Weak Punishment	1.41^{*}	1.39^{*}
	(0.59)	(0.60)
Responder Effort Ratio	0.19	0.06
	(1.01)	(1.03)
Female	-0.44	-0.43
	(0.30)	(0.31)
Age	0.08	0.12
	(0.10)	(0.10)
Native English Speaker	0.65^{*}	0.56^{+}
	(0.29)	(0.30)
Intercept	-2.19	-2.15
	(2.29)	(2.17)
N	99	99
Pseudo R^2	0.12	0.15
χ^2	16.35	19.93

Each column presents random effects probit regressions. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the norms. Individuals who seem to care little for norms may be unaffected by changes in the alternatives.

Table 2.8 regresses the probability that a proposer conditionally offered (8,2) on behavior in the rule following task. The results in column (1) show that proposers who waited longer the lights were more likely to make conditional offers of (8,2). The regression in column (8,2) finds a negative but

Table 2.8: Predicting Conditional Offer of (8,2) by Rule Following

	(1) Conditionally Offer (8,2)	(2) Conditionally Offer (8,2)
Waiting Time	0.03^{+}	
	(0.02)	
Never Waited		−0.44
		(0.48)
Responder Surplus	1.00	1.00
	(0.65)	(0.64)
Weak Punishment	0.37	0.37
	(0.44)	(0.44)
Responder Surplus X Weak Punishment	−0.80	−0.85
	(0.66)	(0.67)
Proposer Effort Ratio	1.08	1.18
	(1.18)	(1.15)
Female	−0.06	0.04
	(0.33)	(0.32)
Age	0.19	0.18
	(0.12)	(0.12)
Native English Speaker	0.18	0.15
	(0.31)	(0.30)
Intercept	-4.99^{+}	−4.12
	(2.84)	(2.71)
N	99	99
Pseudo R^2	0.07	0.06
χ^2	8.19	6.06

Each column presents random effects probit regressions. Standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

insignificant coefficient for proposers who never waited at the lights. Combined with the finding in result 5, it is possible that proposers who care more about norms are more likely to choose responder favorable alternatives .

2.5 Discussion

While this paper has argued in favor of the social norm approach to decision making in bargaining there exist other modeling approaches for concerns for fairness. Models of fairness tend to focus on individual concerns about the distribution of payoffs (Bolton and Okenfels, 2000; Fehr and Schmidt, 1999) or intention-based reciprocity (Dufwenberg and Kirchsteiger, 2004; Rabin, 1993) or a combination of both (Charness and Rabin, 2002; Falk and Fischbacher, 2006; Ridinger, 2015a). These models can account for a number of behavioral regularities that occur in experimental results.

Models of distributional concerns like Bolton and Okenfels (2000) and Fehr and Schmidt (1999) suggest that rejections of the (8,2) offer by responders should not change depending on the alternatives available to the proposer. This is because in these models, what the proposer could have chosen does not influence the perceived fairness of what the person actually chose. The models can account for positive rejections of an (8,2) offer when the alternative is (8,2). That is, when the proposer does not have a choice, outcome-based models can explain rejection by responders.

Intention-based reciprocity models like Rabin (1993) and Dufwenberg and Kirchsteiger (2004) can predict differences between rejections of the (8,2) offer depending on the alternative available. These models are based on psychological game theory (Geanakoplos et al., 1989) which allows beliefs to be included directly into the utility function. What the proposer could have chosen can potentially change individual responder beliefs about the kindness of the proposer. These models can capture the idea that the alternatives available may change the perceived intention of the (8,2) offer. Both the Rabin (1993) and Dufwenberg and Kirchsteiger (2004) models are not able to explain why individuals may still reject an offer of (8,2) when the alternative is (8,2). This is because these models do not account for the idea that outcomes may matter as well.

Models that combine intention-based reciprocity and distributions include Falk and Fischbacher (2006) and Ridinger (2015a). Both of these models can explain differences in rejection rates as the alternatives change and positive rejections of the (8,2) offer when the proposer can only offer (8,2). One drawback of the model in Ridinger (2015a) is that it shares the same functional form as the Dufwenberg and Kirchsteiger (2004) model. Due to the functional form assumptions in Dufwenberg and Kirchsteiger (2004), an offer of (8,2) when the alternative is (2,8) is viewed as more unkind compared to when the alternative is (5,5). Both models predict that there should be higher rejection rates of (8,2) when the alternative is (2,8) compared to when the alternative is (5,5). The model in Falk and Fischbacher (2006) does a better job at qualitatively predicting the observed differences in rejection rates of (8,2) as the alternatives change found in Falk et al. (2003) and in the above experiment.

Both the outcome and intention based models do predict that rejection by responders should decrease as the strength of punishment by the responder decreases. In the outcome based models, a decrease in the effectiveness of punishment changes the difference in payoffs between the proposer and responder in the case of rejection. As a result, weak punishment is less attractive compared to when punishment makes both parties equally worse off. In the intention based models, equilibrium exist where rejection is less likely to occur because of reciprocity. In these models, people punish because the act of punishment directly increases their utility. Lowering the effectiveness of the punishment leads to lower gains in utility by punishing.

While these fairness models can account for a number of behavioral regularities that occur in the laboratory, they are unable to explain why prior ownership influenced individual behavior. In addition, the models have difficulty explaining a number of other experimental results (Charness, 2004; Charness and Levine, 2007; Falk et al., 2008, 2003; Falk and Kosfeld, 2006; McCabe et al., 2000, 2003; **?**). One reason for this is that the models are accounting only for a few specific fairness concerns, and subsequently lack the flexibility to explain other relevant factors in decision making.

It is important to note that changes in prior ownership or real effort production should have no influence on rejection rates in any of the distribution, intention-based reciprocity, or combined models. The reason is that these models do not account for different situations or norms that can change how people view the exact same game. When the responder earned the right they may have felt entitled to a much larger share of the surplus compared to when the surplus is just given to subjects by the experimenter. When the responder earned the right, high rejection rates of (8,2) when the proposer could have chosen (10,0) could have occurred due to what responders felt they deserved based on the norm.

2.6 Conclusion

The results from the experiment suggest that prior ownership and the strength of punishment are important in explaining differences in proposer and responder behavior in joint production. When their ability to punish was weak, responders were less likely to reject unfair offers when the proposer had the right compared to responders had the right. When their ability to punish was strong, responders rejected at high rates regardless of who received the right. Proposers were less likely to offer unequal divisions when the responder had the right and had a strong ability to punish. However, when punishment was weak responders did not appear to respect prior ownership by responders.

These findings suggest that both the strength of enforcement and prior ownership in bargaining can influence the degree of conflict in negotiations. Additionally, the treatment variables changed the distribution of subjects who conditionally accepted or rejected the (8,2) offer. This suggests that both prior ownership and punishment changed how individuals viewed the offers of the proposer. These results cast doubt on game theoretic models of fairness that do not account for factors outside of the game that may change individual behavior. Alternative models that rely on social norms may provide a way to model the observed behavior in the experiment (Bicchieri, 2006), because the relevant norms can differ depending on the context.

Chapter 3

Emotions, rule-following, and bargaining norms

3.1 Introduction

Everyday life is filled with social norms that influence individual decision making. Social norms have been important in understanding a wide range of human behavior including: cooperation in collective action problems (Ostrom, 2000), helping behavior in the workplace (Grant and Patil, 2012) , worker lateness (Blau, 1995), and tipping (Azar, 2007). Theoretical models of social norm adherence often posit that an individual's utility is directly influenced by norms (Bicchieri, 2006; Lopez-Perez, 2008; Kessler and Leider, 2012; Kimbrough and Vostroknutov, 2015). Crucial to these models is the assumption that when faced with the same decision problem, individual actions can differ depending on individual differences in the propensity to follow norms. This difference is what can explain why some individuals choose to follow the norm while others do not. The propensity to follow social norms is an important part of the theoretical models, however, less is known about what individual differences may influence this propensity.

One potential source for individual heterogeneity in norm adherence may have to do with moral emotions. Two powerful moral emotions are guilt and shame. Both guilt and shame are negative affective states that may activate in individuals when they transgress norms. To avoid these negative feelings people choose actions that are consistent with socially acceptable behavior. Understanding what others will do in game-theoretic situations depends crucially on the ability to recognize understand the beliefs of others and predict their behavior. This is potentially another source for individual differences in social norm adherence. The ability to understand others emotions, thoughts, and beliefs is known as theory of mind (ToM). Higher ToM may make individuals more likely to recognize how others would feel about their actions (Singer and Fehr, 2005). This recognition could lead higher ToM types to be less likely to transgress norms.

In this paper: I ask: How do differences in both the propensity to feel moral emotions and the ability to recognize the emotions of others impact whether individuals follow rules and punish others for violating these rules? I address this question by examining how individual emotions and emotional recognition influence individual decisions in two different environments. Specifically, individual decisions to abide by bargaining norms and follow costly rules. Using psychometric measures of individual propensities to feel shame and guilt, I examine how subject differences in these moral emotions influence behavior in these two settings. Using a measure of ToM, this paper explores how differences in the ability to recognize the emotions of others influences both rule-following and bargaining behavior.

Using psychometric measures of guilt, shame, and ToM, I explore these theories using data from an ultimatum game experiment examining prior ownership norms (Ridinger, 2015b). Guilt was not predictive of behavior; however, both shame and ToM were. Individuals who had greater feelings of shame were much more likely to respect prior ownership of responders and punish proposers who transgressed prior ownership of responders. Individuals who scored higher in ToM were much more likely to respect prior ownership by proposers. In addition, I find similar result using a more direct test of the effects of guilt, shame, and ToM on norm adherence. Individuals participated in a rule-following task designed to capture the propensity to follow social norms. The results show that individuals who feel more shame and have higher ToM are more likely to follow the rules.

Bargaining situations often are strongly influenced by social norms (Elster, 1989; Joseph Henrich and McElreath, 2001; Bicchieri, 2006; Kessler and Leider, 2012). One particularly powerful norm in bargaining is prior ownership (Leliveld et al., 2008; Ridinger, 2015b). In a recent experiment, Ridinger (2015b) showed that prior ownership had a strong effect on behavior in ultimatum games with jointly produced surpluses. The results from the experiment suggest that many subjects respected prior ownership and punished those who transgressed these ownership norms. While the results from the experiment suggest that prior ownership is an important social norm in bargaining, it is unclear what individual factors influence individual responses to these norms.

Guilt and shame are negative emotions that humans often experience. Shame and guilt are typically used as synonyms in both everyday life and academic research. However, researchers have argued that that there exists key differences between the two affective states (Lewis, 1971; Tagney et al., 1996; Tracy and Robins, 2004; Tagney et al., 2007; Cohen et al., 2011). One approach has centered on the distinction between the self and behavior (Lewis, 1971; Tracy and Robins, 2004). Individuals experiencing guilt often focus on a specific behavior ("I did something bad"), but when experiencing shame focus on their individual self ("I am a bad person"). A separate approach attempts to distinguish guilt and shame based on whether a transgression is observed by others or not. This public versus private distinction suggests that guilt is activated when a transgression is private whereas shame is activated when others are aware of the transgression(Lewis, 1971; Tracy and Robins, 2004). Recent research suggests that these different distinctions are complementary (Wolf et al., 2010; Cohen et al., 2011). (Wolf et al., 2010; Cohen et al., 2011) present evidence that shame tends to be strongly related to situations that are public and that when experiencing shame people tend to make negative judgments about their self. Similarly, guilt appears to be strongly correlated with situations that are private and when experience guilt people tend to make negative judgments about their behavior.

Feelings of guilt and shame are likely to be activated when a person violates a rule or norm. Individuals may attempt to avoid actions that may make them feel guilt or shame. Game-theoretic models using psychological game theory have attempted to capture feelings of guilt by proposing that people may be guilt averse (Battigalli and Dufwenberg, 2009; Charness and Dufwenberg, 2006). Individuals try to avoid decisions that they believe will make them feel guilty. While this avoidance may influence the decision to adhere to a norm, guilt and shame may also be related to how people judge the guilt and shame they believe others should feel for violating a norm.

In order to measure guilt and shame, this paper uses uses the guilt and shame proneness scale (GASP) (Cohen et al., 2011). The scale has been used in a number of studies to measure individual differences in the propensity to feel guilt and shame. Howell et al. (2012) used the GASP and found that individuals who had higher guilt proneness were more likely to report a general willingness to apologize to others. No correlation was found for shame proneness as measured by the GASP. Bracht and Regner (2013) find that individuals who scored higher in guilt proneness where more pro-social in the mini-trust game. Jordan et al. (2015) found that guilt proneness was not related to self-forgiveness, but positively related with forgiving others. In addition, shame proneness was negatively related to forgiving oneself and others. While Carpenter et al. (2016) found that guilt-proneness was positively associated with forgiving oneself while shame-proneness was negatively associated with self-forgiveness. These findings may be a potential reason why shame avoidance may be powerful in public settings as people who feel great shame are less likely to forgive themselves for transgressions. Ent and Baumeister (2015) found that people higher in guilt proneness, as measured by the GASP, valued harm avoidance more than obedience to authority and were more likely to disobey the experimenter to alleviate suffering of another individual. Arli et al. (2016) find that people who feel higher guilt and shame are less likely to report that they engage in unethical consumer behavior. The effect of shame is strongest in Australia, the authors argue that this may be due to Australians individualistic culture compared to Indonesia.

While individual differences in the propensity to feel moral emotions may influence behavior, in game-theoretic situations it is also important to predict the behavior of others. The ability of understand what others will do in game-theoretic situations depends crucially on the ability to recognize the others utility function and their beliefs. This ability known as ToM and often develops in young children (Baron-Cohen et al., 2001). Adults can have varying levels of ToM ability. For example, individuals on the autism spectrum often exhibit lower ToM ability and females typically manifest slightly higher ToM ability compared to males (Kirkland et al., 2013).

Previous research has suggest that ToM helps individuals recognize gains from cooperation and predict the behavior of others (McCabe et al., 2000; Singer and Fehr, 2005). Research on children and adolescents has found differences in judgments about the social appropriateness of actions between those with and without autism (Loveland et al., 2001). Using functional magnetic resonance imaging (FMRI), Berthoz et al. (2002) found increased activation in areas in the brain thought to be responsible for ToM when subjects read stories about norm violations. Higher ToM ability may make individuals more likely to recognize how others would feel about their actions (Singer and Fehr, 2005). This recognition could lead higher ToM types to be less likely to transgress norms.

To measure affective ToM ability, this paper uses the reading the mind in the eyes test (RMET) (Baron-Cohen et al., 2001). The RMET has been used in numerous studies to measure individual differences in ToM ability in the adult population. Researchers examining the relationship between RMET and individual behavior have examined trading markets (Bruguier et al., 2010; Martino et al., 2013), incentives (Ridinger and McBride, 2015), and strategic sophistication (Georganas et al., 2015).

3.2 Experiment

This paper uses data collected in Ridinger (2015b) where a total of 198 students from a large public university participated in an experiment at a computer laboratory. Subjects were recruited from a large and diverse subject pool. Prior to each experimental session, a randomly selected group of students from the subject pool were sent an email about participating in the upcoming session. Students then registered through the subject pool website. At the day of the experiment, registered students were randomly assigned to a computer terminal. No subject participated in more than one experimental session. To facilitate data collection, the experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

3.2.1 Guilt and Shame Measure

To measure feelings of guilt and shame, this paper uses the GASP scale (Cohen et al., 2011). The GASP consists of a series of statements where individuals indicate on a seven point likert scale the likelihood that they would react in the way described. The four statements in the guilt judgment sub-scale attempt to measure the degree in which individuals feel guilty about private transgressions. An example of one of the guilt statements is "After realizing you have received too much change at a store, you decide to keep it because the salesclerk doesn't notice. What is the likelihood that you would feel uncomfortable about keeping the money?" The four statements in the shame sub-scale are designed to assess how much individuals feel shame about public transgressions. An example of one of the shame behavior statements is "You give a bad presentation at work. Afterwards your boss tells your coworkers it was your fault that your company lost the contract. What is the likelihood that you would feel incompetent?" This public versus private distinction is an important difference between whether someone is experiencing feelings of guilt or shame. While guilt and shame may be different moral emotions, they are nevertheless very similar and the two sub-scales have been shown to be highly correlated (Cohen et al., 2011).

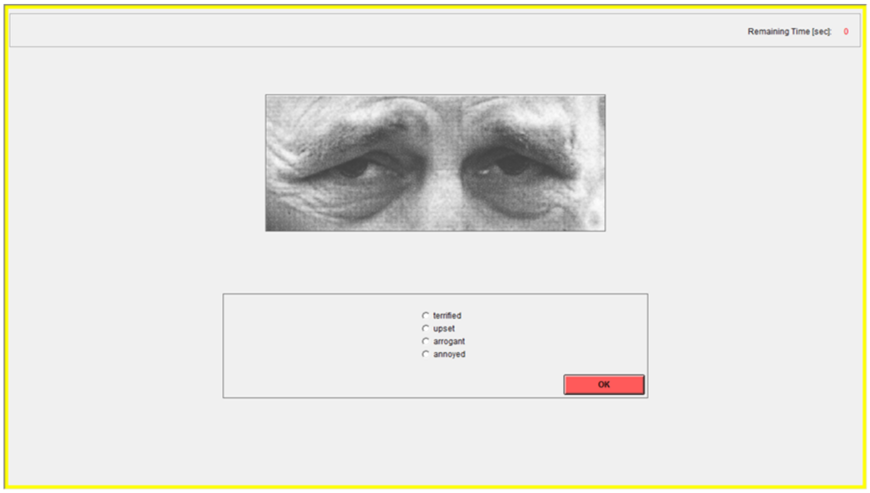


Figure 3.1: Example of RMET Question

3.2.2 Affective ToM Measure

To measure affective ToM subjects completed the RMET (Baron-Cohen et al., 2001). The RMET consists of a series of 36 individual pictures of the area around a person's. Each photograph is accompanied by four words (see Figure 1). Subjects chose the word that best describes what the person is thinking or feeling. Each question was answered without feedback. Subjects were provided with a printed handout of word definitions for all the words used in the task. The RMET has been shown to be a consistent measure of affective ToM ability and used in wide range of studies (Baron-Cohen et al., 2001; Golan et al., 2006; Torralva et al., 2007; Kirkland et al., 2013; Ridinger and McBride, 2015).

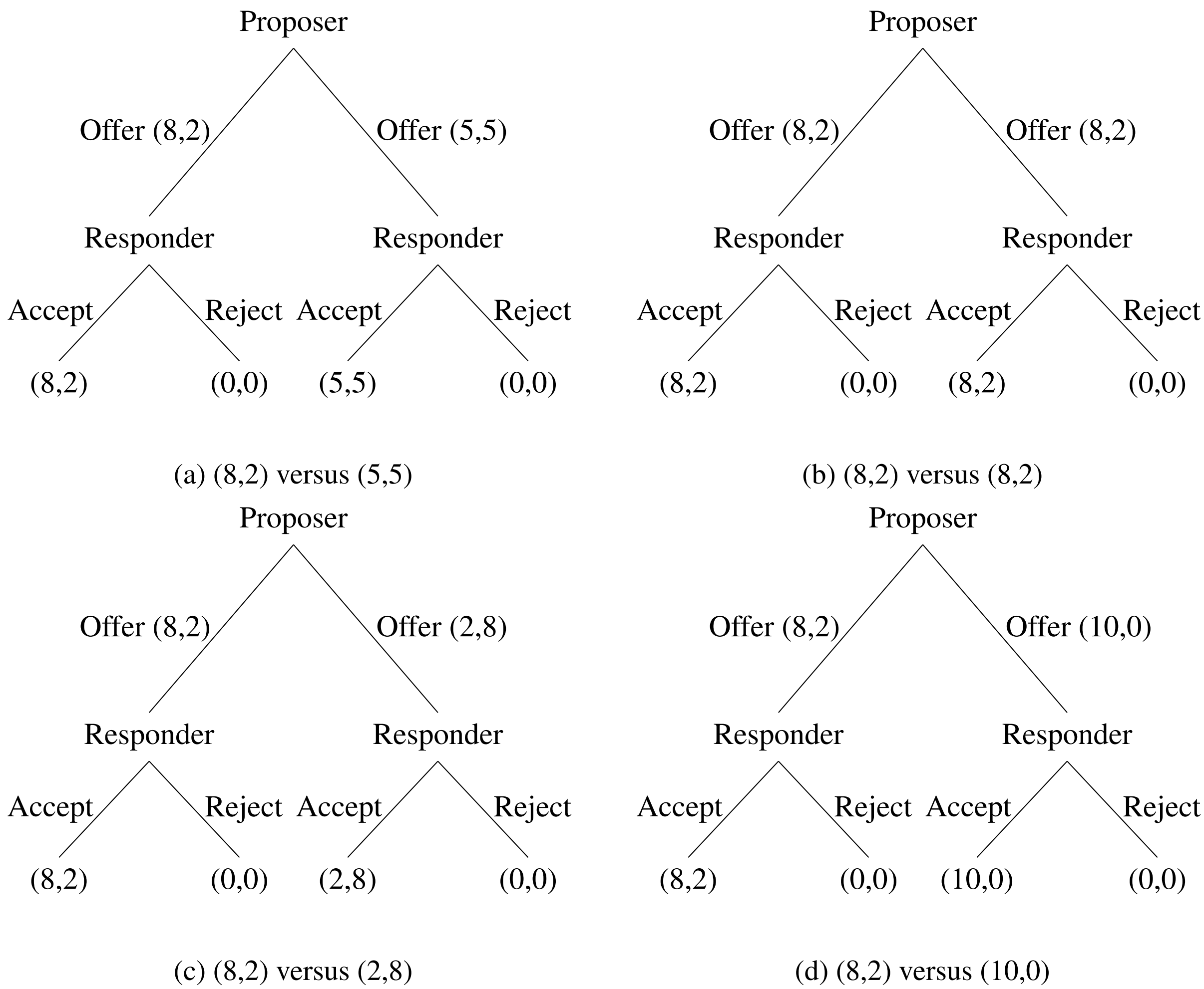
3.2.3 Bargaining

Ridinger (2015b) randomly paired subjects who then participated in the slider task (Gill and Prowse, 2012). The slider task is a real-effort task where each subject must move sliders across the screen. The task is difficult for a subject to finish before the time runs out. The advantage of using the slider task is that the task is not as cognitively difficult as other real-effort task like knowledge based questions or mathematical computations. This may reduce potential confounds in observed effort that are due to differences in individual knowledge or mathematical ability as opposed to actual effort. Subject completed the task three times with the first two rounds serving as practice rounds.

Their combined efforts from the real-effort task jointly produced a surplus. The entire surplus was then allocated to the subject who expended higher effort. After learning the outcome, the two subjects played a series of four mini-ultimatum games on how to divide the surplus they created in the real effort task (see Figure 2). In each min-ultimatum game, proposers could select between an offer of $8 for herself and $2 denoted (8,2) and an alternative offer which varied across the games. The outcome from one of the four mini-ultimatum games was randomly chosen for payment. Experimental conditions varied whether the person who received the surplus was the proposer (Proposer Surplus treatment) or the responder (Responder Surplus treatment). In addition, treatments varied the strength of punishment by the responder. In the Strong Punishment treatment, rejection by the responder left both subjects with $0. In the Weak Punishment treatment, rejection by the responder left the proposer with $2 and the responder with $0.

3.2.4 Rule-Following

In Ridinger (2015b), each subject participated in the rule-following task designed to measure individual sensitivity to social norms or rule-following (Kimbrough and Vostroknutov, 2015) . In the task, each subject controlled a stick figure that walked from the left side to the right side of the screen. As the figures walked, they automatically stopped at a traffic light. There were a total of five traffic lights. At any time, subjects could start walking again by pressing the walk button. After five seconds the light would turn green. Subjects began with $7 and were told that for each second they spent on the task they would loose $0.07. Walking without waiting at any of the lights took a total of 24 seconds. Subjects who did not wait would loose at least $1.68. Subjects who waited the full five seconds at each light lost $3.08. In the instructions, subjects were told “The



(a) (8,2) versus (5,5)

(b) (8,2) versus (8,2)

(c) (8,2) versus (2,8)

(d) (8,2) versus (10,0)

Figure 3.2: Mini-Ultimatum Games with Strong Punishment

rule is to wait at each stop light until it turns green." An example of what a typical subject would see in the task can be seen in figure 1.

In essence, the rule-following task created a situation where a person could follow a rule where adherence to the rule is costly to themselves. The time that subjects took to complete the task can be viewed as a measure of an individuals sensitivity to following rules or social norms. Kimbrough and Vostroknutov (2015) show that this measure is good predictor of subject behavior in wide range of experimental games. In addition, they show that the rule that subjects are told appears to be what is leading some subjects to wait. Without the rule, the majority of subjects ignore the rule, suggesting that the difference in behavior was driven by the stated rule.

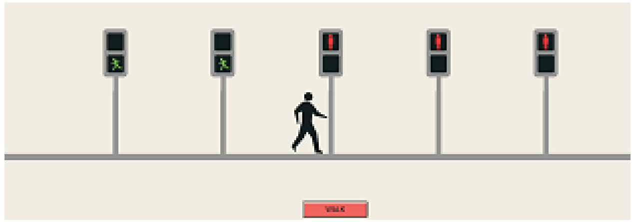


Figure 3.3: Example of Rule-Following Task

3.2.5 Descriptive Statistics

Table 3.1 presents the summary statistics for the psychometric scales and demographic information. Both Shame and Guilt were generated by adding subjects scores from the GASP sub scale. Larger numbers indicated greater shame and guilt proneness. The RMET is equal to the number of questions where subjects selected the correct word that best describes what the person in the image was thinking or feeling. Waiting Time is equal to the total time that subjects waited at each light. Overall, 66% of the subjects were female.

Table 3.2 shows the summary statistics for the bargaining part of the experiment. Overall rejection and offer rates of the (8,2) offer are displayed and project down by alternative. Each session took approximately one hour to complete and subjects took home an average of $14.47.

One potential worry is that there may be group differences in scores of Shame, Guilt, or RMET across the treatments. If this is occurred, then we could falsely conclude that differences in these psychometric measures are explaining treatment differences when in fact it is due to randomness. In table C.8 in Appendix C, non-parametric tests show that we fail to reject the null hypotheses that the scores for Shame, Guilt, and RMET come from the same distribution across the treatments. This suggests that the scores for Shame, Guilt, and RMET are similar across the treatments and makes it less likely to be a potential confound. Histograms for Shame, Guilt, and RMET can be found in Appendix C.

Table 3.1: Summary Statistics

	Mean	Standard Deviation
Shame	22.21	4.31
Guilt	20.70	4.73
RMET	27.09	4.05
Waiting Time	21.95	7.80
Female	0.66	0.47
Age	20.37	1.54
Native English Speaker	0.62	0.49
Observations	198	

Table 3.2: Treatment Summary Statistics

	Strong Punishment		Weak Punishment		
	Proposer Right	Responder Right	Proposer Right	Responder Right	All Treatments
Overall Reject (8,2)	0.44	0.47	0.14	0.33	0.35
Reject (8,2) vs Alt (5,5)	0.63	0.64	0.22	0.46	0.49
Reject (8,2) vs Alt (2,8)	0.48	0.44	0.13	0.29	0.34
Reject (8,2) vs Alt (8,2)	0.44	0.44	0.09	0.33	0.33
Reject (8,2) vs Alt (10,0)	0.22	0.36	0.13	0.25	0.24
Overall Offer (8,2)	0.78	0.63	0.74	0.73	0.72
Offer (8,2) vs Alt (5,5)	0.41	0.24	0.48	0.33	0.36
Offer (8,2) vs Alt (2,8)	0.89	0.64	1.00	0.92	0.86
Offer (8,2) vs Alt (10,0)	0.89	0.84	0.74	0.88	0.84
Take Home Earnings	14.11	14.40	15.03	14.41	14.47
Observations	54	50	46	48	198

3.3 Predictions

Ridinger (2015b) suggests that the allocating the surplus prior to bargaining created a sense of ownership. Subjects who received the right may have felt that prior ownership entitled them to a greater share of the surplus in the bargaining phase of the experiment. Subjects who did not receive the surplus prior may have respected the prior ownership of others.

Hypothesis 1: *(a) Responders who have a higher propensity to feel guilt, should be less likely to reject unequal offers in the Proposer Surplus treatment compared to the Responder Surplus treatment. (b) Proposers who have a higher propensity to feel guilt, should be more likely to propose unequal offers in the Proposer surplus treatment compared to the Responder Surplus treatment.*

Hypothesis 2: *(a) Responders who have a higher propensity to feel shame, should be less likely to reject unequal offers in the Proposer Surplus treatment compared to the Responder Surplus treatment. (b) Proposers who have a higher propensity to feel shame, should be more likely to propose unequal offers in the Proposer surplus treatment compared to the Responder Surplus treatment.*

Hypotheses 1 and 2 come from the argument that both shame and guilt create a negative state in individuals. To avoid this negative feeling, individuals may choose actions that are less likely to make them feel shame or guilt. In addition, higher propensity to feel guilt or shame may also make it more likely that people will be upset at the actions of others if they deviate from accepted social behavior. For example, responders who have higher propensity to feel shame and received the surplus prior to bargaining, may be even more likely to reject unequal offers than those lower in shame. This could occur if shame proneness is highly correlated with not only the shame a person feels about their own actions but also how much shame they believe others should feel about their actions.

Hypothesis 3: *(a) Responders who have a higher affective ToM ability, should be less likely to reject unequal offers in the Proposer Surplus treatment compared to the Responder Surplus treatment. (b) Proposers who have a higher affective ToM ability, should be more likely to propose unequal offers in the Proposer surplus treatment compared to the Responder Surplus treatment.*

Hypothesis 3 relies crucially on individuals being able to recognize what others believe the norm should be. Ridinger (2015b) argues that the Proposer and Responder Surplus treatments activate different norms about what individuals should receive. When a proposer had the surplus prior to bargaining, a responder with high ToM ability may be more likely to recognize that the proposer expects to receive a greater share of the surplus. As a result, responders with high ToM ability may be less likely to reject unequal offers in the Proposer Surplus treatment compared to the Responder Surplus. Similarly, proposers with high ToM ability may be likely to realize that responders who had the surplus prior to bargaining will expect to receive a higher share. Due to this recognition, proposers with high ToM ability could be less likely to propose unequal offers in the Responder Surplus treatment compared to the Proposer Surplus treatment.

Hypothesis 4: *Higher affective ToM ability and greater propensity to feel guilt and shame are more likely to follow the rule in the rule-following task.*

Hypothesis 5: *Since actions are observable in both the bargaining experiment and rule-following task, the effects of shame should be a stronger influence on behavior compared to guilt.*

Hypothesis 5 comes directly from previous research that argues that shame is more likely to be activated in public situations whereas guilt is more likely to be activated in private situations (Lewis, 1971; Tracy and Robins, 2004; Cohen et al., 2011). In the bargaining experiment, the actions of both the proposer and responder are observable. The decisions of the proposer and responder are known both to each other and to the experimenter. In the rule-following task, the actions of the individuals to wait at the light are not observable to the other subjects. However, the actions that the individuals make are observable to the experimenter. If subjects think of their actions as observable to the experimenter, then shame may be a stronger influence on behavior. While both guilt and shame may be activated both of these situations, it is possible that the effect of shame will be stronger than guilt.

3.4 Results

3.4.1 Bargaining Results

Result 1: *Responders with greater shame proneness were more likely to reject unequal offers in the Responder Surplus treatment. Proposers with higher shame proneness were less likely to propose unequal offers.*

The results in table 3.3 come from probit regressions predicting the probability that the responder will reject the (8,2) offer by differences in shame. From column (1), it is clear that responders who report higher sensitivity to feelings of shame are significantly more like to reject the unequal offer. Column (2) shows that this is driven through the effect of shame in the Responder Surplus treatment. Individuals with higher scores on the shame scale are significantly more likely to reject the (8,2) offer when they received the surplus prior to bargaining. Column (3) shows that differences in shame do not seem to deferentially impact rejection rates as the strength of punishment varies. The effect of shame in the Responder Surplus treatment remains even when controlling for the interaction between shame and the Weak Punishment treatment. Using the probit specification from column (2), figure 3 presents the predicted probability of rejecting an (8,2) offer at different levels of shame. Individuals who have high level of propensity to feel shame reject (8,2) offers more often in the Responder Surplus compared to the Proposer Surplus. At lower levels of shame, this is reversed. This suggests that individuals who feel low shame may be willing to accept unequal offers when they had prior ownership.

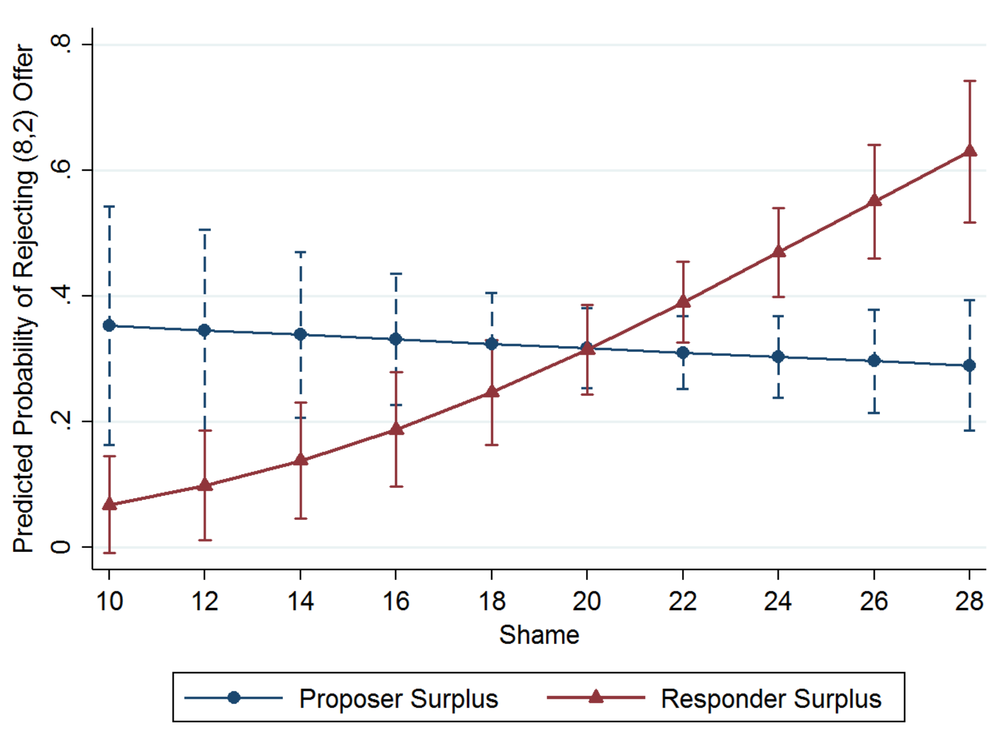


Figure 3.4: Predicted Probability of Rejecting (8,2) Offer by Shame with 95% Confidence Intervals

Table 3.4 presents probit regressions predicting the probability that the proposer will offer (8,2) by differences in shame. The coefficient for shame is negative a significant in columns (1) and (3). This suggests that overall individuals who reported greater likelihood of feeling shame were less likely to offer (8,2) when there was a more responder favorable alternative available. Interestingly, columns (2) and (4) show that there was not a differential impact of shame in the Responder Surplus treatment compared to the Proposer Surplus treatment. The significant effect of shame is no longer significant, although the coefficients remain negative, suggesting that the interaction with Responder Surplus is introducing multicollinearity.

Result 2: *No significant correlation was found between responder or proposer behavior and guilt proneness.*

The results in table 3.5 show that differences in guilt are not predictive of rejections of unequal offers. The coefficient for guilt is insignificant across specifications. Using the probit specification

Table 3.3: Predicting Rejection of (8,2) Offer by Differences in Shame

	(1) Reject (8,2)	(2) Reject (8,2)	(3) Reject (8,2)	(4) Reject (8,2)
Shame	0.05*	−0.01	0.07**	0.01
	(0.03)	(0.04)	(0.03)	(0.04)
Responder Surplus	0.61	−2.33*	0.65	−2.21*
	(0.51)	(1.26)	(0.52)	(1.25)
Weak Punishment	−0.25	−0.28	1.50	1.36
	(0.40)	(0.40)	(1.22)	(1.19)
Shame X Responder Surplus		0.13**		0.13**
		(0.05)		(0.05)
Shame X Weak Punishment			−0.08	−0.07
			(0.05)	(0.05)
Effort Ratio	1.13	1.18	1.14	1.20
	(0.82)	(0.84)	(0.82)	(0.85)
Alternative (5,5)	0.76***	0.79***	0.77***	0.80***
	(0.16)	(0.16)	(0.16)	(0.16)
Alternative (2,8)	0.31**	0.34**	0.32**	0.34**
	(0.15)	(0.15)	(0.15)	(0.15)
Alternative (8,2)	0.28**	0.30**	0.29**	0.31**
	(0.13)	(0.13)	(0.13)	(0.13)
Female	−0.37	−0.46*	−0.35	−0.44*
	(0.25)	(0.25)	(0.25)	(0.25)
Intercept	−5.49***	−3.59*	−6.49***	−4.62**
	(1.95)	(2.02)	(2.09)	(2.25)
N	392	392	392	392
pseudo R^2	0.14	0.16	0.15	0.17

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Predicting Offer of (8,2) by Differences in Shame

	(1) Offer (8,2)	(2) Offer (8,2)	(3) Offer (8,2)	(4) Offer (8,2)
Shame	-0.07^{**}	-0.05	-0.07^{*}	-0.04
	(0.03)	(0.04)	(0.04)	(0.05)
Responder Surplus	-0.54	0.16	-0.53	0.22
	(0.44)	(1.31)	(0.45)	(1.37)
Weak Punishment	0.72	0.73	1.12	1.21
	(0.48)	(0.48)	(1.49)	(1.52)
Shame X Responder Surplus		-0.03		-0.04
		(0.05)		(0.06)
Shame X Weak Punishment			-0.02	-0.02
			(0.06)	(0.06)
Effort Ratio	-0.20	-0.30	-0.18	-0.29
	(1.01)	(1.00)	(1.01)	(1.00)
Alternative (5,5)	-1.68^{***}	-1.69^{***}	-1.68^{***}	-1.69^{***}
	(0.20)	(0.20)	(0.19)	(0.20)
Female	0.19	0.21	0.20	0.22
	(0.25)	(0.25)	(0.25)	(0.25)
Intercept	4.26^{***}	3.90^{**}	4.12^{**}	3.71^{**}
	(1.60)	(1.72)	(1.66)	(1.82)
N	194	194	194	194
pseudo R^2	0.31	0.31	0.31	0.31

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Predicting Rejection of (8,2) Offer by Differences in Guilt

	(1) Reject (8,2)	(2) Reject (8,2)	(3) Reject (8,2)	(4) Reject (8,2)
Guilt	−0.00	−0.01	0.02	0.01
	(0.02)	(0.04)	(0.03)	(0.04)
Responder Surplus	0.69	0.25	0.70	0.09
	(0.51)	(1.02)	(0.54)	(1.01)
Weak Punishment	−0.21	−0.21	1.13	1.21
	(0.41)	(0.41)	(1.10)	(1.17)
Guilt X Responder Surplus		0.02		0.03
		(0.05)		(0.05)
Guilt X Weak Punishment			−0.06	−0.07
			(0.05)	(0.05)
Effort Ratio	0.83	0.81	0.81	0.79
	(0.84)	(0.84)	(0.85)	(0.86)
Alternative (5,5)	0.75***	0.75***	0.76***	0.76***
	(0.15)	(0.15)	(0.15)	(0.15)
Alternative (2,8)	0.31**	0.31**	0.31**	0.31**
	(0.14)	(0.14)	(0.14)	(0.14)
Alternative (8,2)	0.28**	0.28**	0.28**	0.28**
	(0.12)	(0.12)	(0.12)	(0.12)
Female	−0.24	−0.26	−0.21	−0.23
	(0.23)	(0.23)	(0.25)	(0.24)
Intercept	−3.90*	−3.67*	−4.27**	−3.99*
	(2.02)	(2.01)	(2.14)	(2.10)
N	392	392	392	392
pseudo R^2	0.12	0.12	0.13	0.13

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Predicting Offer of (8,2) by Differences in Guilt

	(1) Offer (8,2)	(2) Offer (8,2)	(3) Offer (8,2)	(4) Offer (8,2)
Guilt	−0.02	0.00	−0.02	0.01
	(0.02)	(0.03)	(0.03)	(0.04)
Responder Surplus	−0.62	0.37	−0.63	0.36
	(0.43)	(1.13)	(0.42)	(1.13)
Weak Punishment	0.74*	0.72	0.96	0.96
	(0.45)	(0.46)	(1.17)	(1.19)
Guilt X Responder Surplus		−0.05		−0.05
		(0.05)		(0.05)
Guilt X Weak Punishment			−0.01	−0.01
			(0.05)	(0.05)
Effort Ratio	−0.16	−0.33	−0.17	−0.34
	(1.00)	(1.00)	(1.00)	(0.99)
Alternative (5,5)	−1.61***	−1.62***	−1.61***	−1.62***
	(0.18)	(0.18)	(0.18)	(0.18)
Female	0.13	0.15	0.13	0.15
	(0.25)	(0.25)	(0.25)	(0.25)
Intercept	2.86*	2.51	2.79*	2.44
	(1.56)	(1.59)	(1.63)	(1.67)
N	196	196	196	196
pseudo R^2	0.29	0.29	0.29	0.29

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

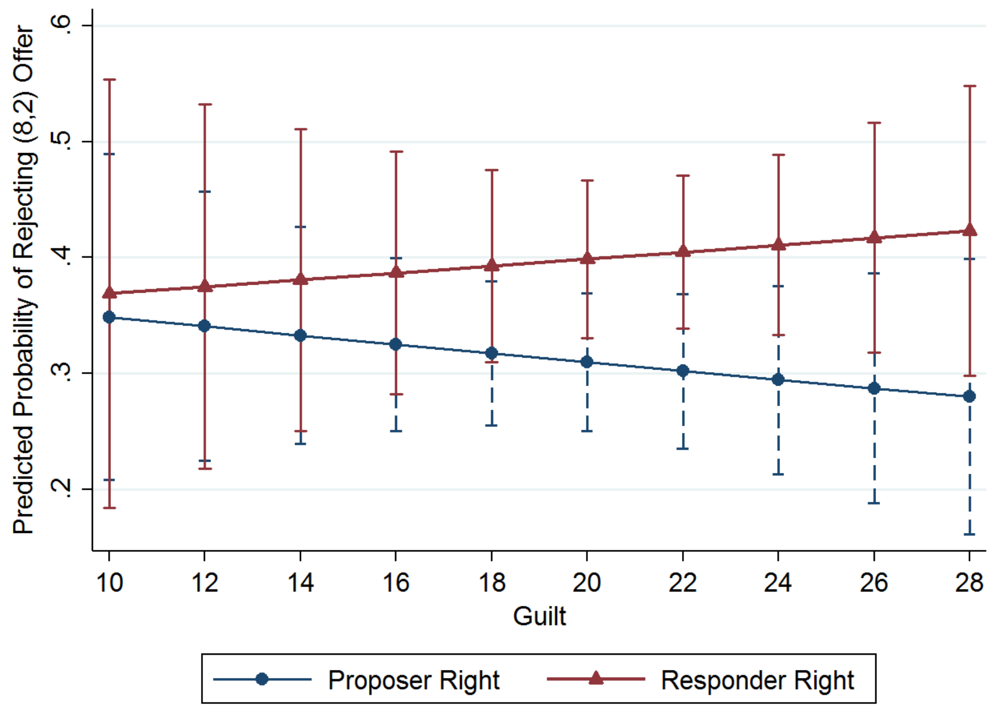


Figure 3.5: Predicted Probability of Rejecting (8,2) Offer by Guilt with 95% Confidence Intervals

from column (2), figure 5 presents the predicted probability of rejecting an (8,2) offer at different levels of guilt. The change in slopes for Proposer and Responder Surplus are consistent with the predictions, but it is clear that the slope is small. In addition, the confidence intervals for the estimates are quite large. A potential explanation for this result is that the guilt scale measures feelings from private transgressions. However, in this experiment actions by both the proposer and the responder where known to each other. As a result, it is more likely that the shame scale that measures feelings from public transgressions would be more predictive of subject behavior

Similar results for proposers can be found in table 3.6. Higher likelihood to report feeling guilt does not significantly predict unequal offers by proposers. This result holds across specifications.

Result 3: *Responders with higher affective ToM ability were less likely to reject unequal offers in both treatment conditions. Proposers with higher affective ToM ability were less likely to propose unequal offers in the Weak Punishment treatment compared to the Strong Punishment treatment.*

Table 3.7: Predicting Rejection of (8,2) Offer by Differences in RMET

	(1) Reject (8,2)	(2) Reject (8,2)	(3) Reject (8,2)	(4) Reject (8,2)
RMET	-0.04	-0.08^{**}	-0.03	-0.06
	(0.03)	(0.04)	(0.04)	(0.04)
Responder Surplus	0.57	-1.22	0.56	-1.56
	(0.50)	(1.69)	(0.51)	(1.77)
Weak Punishment	-0.41	-0.25	0.15	0.92
	(0.43)	(0.44)	(1.54)	(1.74)
RMET X Responder Surplus		0.07		0.08
		(0.06)		(0.06)
RMET X Weak Punishment			-0.02	-0.04
			(0.05)	(0.06)
Effort Ratio	0.81	1.00	0.82	1.05
	(0.84)	(0.87)	(0.84)	(0.87)
Alternative (5,5)	0.75^{***}	0.75^{***}	0.75^{***}	0.76^{***}
	(0.15)	(0.15)	(0.15)	(0.15)
Alternative (2,8)	0.31^{**}	0.31^{**}	0.31^{**}	0.31^{**}
	(0.14)	(0.14)	(0.14)	(0.14)
Alternative (8,2)	0.28^{**}	0.28^{**}	0.28^{**}	0.28^{**}
	(0.12)	(0.13)	(0.12)	(0.13)
Female	-0.25	-0.30	-0.26	-0.32
	(0.23)	(0.24)	(0.23)	(0.24)
Intercept	-1.16	-0.35	-1.46	-0.84
	(2.11)	(2.17)	(2.27)	(2.21)
N	396	396	396	396
pseudo R^2	0.12	0.13	0.12	0.13

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Predicting Offer of (8,2) by Differences in RMET

	(1) Offer (8,2)	(2) Offer (8,2)	(3) Offer (8,2)	(4) Offer (8,2)
RMET	−0.00	−0.00	0.07	0.06
	(0.03)	(0.03)	(0.05)	(0.05)
Responder Surplus	−0.59	−0.69	−0.68	−1.04
	(0.42)	(1.58)	(0.42)	(1.63)
Weak Punishment	0.76*	0.77*	4.14**	4.18**
	(0.45)	(0.45)	(1.71)	(1.76)
RMET X Responder Surplus		0.00		0.01
		(0.06)		(0.06)
RMET X Weak Punishment			−0.12**	−0.12**
			(0.06)	(0.06)
Effort Ratio	−0.11	−0.11	0.02	0.03
	(1.00)	(1.00)	(1.01)	(1.01)
Alternative (5,5)	−1.61***	−1.61***	−1.65***	−1.66***
	(0.19)	(0.18)	(0.19)	(0.19)
Female	0.09	0.09	0.04	0.05
	(0.25)	(0.25)	(0.25)	(0.25)
Intercept	2.53	2.58	1.47	1.62
	(1.69)	(1.73)	(1.81)	(1.86)
N	198	198	198	198
pseudo R^2	0.29	0.29	0.31	0.31

All regressions are probit with clustered standard errors at the subject level in parentheses.
Additional controls include session fixed effects, if English is the native language, and age.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

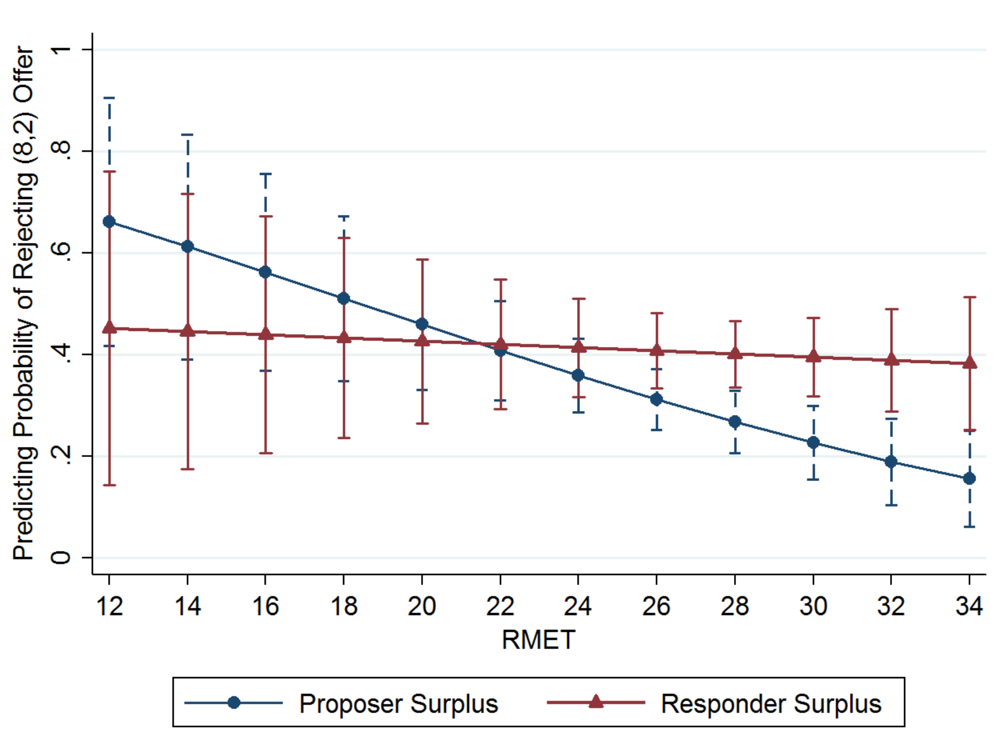


Figure 3.6: Predicted Probability of Rejecting (8,2) Offer by RMET with 95% Confidence Intervals

Table 3.7 shows responders who scored higher on the RMET where less likely to reject unequal offers when the proposer had prior ownership. The RMET attempts to measure subjects ability to recognize the emotions of others. The ability to recognize the emotions of others may have led responders to take into account how their actions would affect the proposer. This in turn led them to be more likely to respect the prior ownership of others.

Table 3.8 examines how theory of mind influenced proposer offers. Column (1) shows that the coefficient for RMET does not seem to predict (8,2) offers. Interacting RMET with a dummy variable for the Responder Surplus treatment shows no significant relationship. It appears that differences in affective ToM between proposers may not lead to differential offers due to changes in ownership. However, columns (3) and (4) show that when higher RMET is interacted with a dummy variable for the Weak Punishment treatment, individuals with higher RMET were less likely to proposer an unequal offer when responder punishment is weak. Despite responders weakened bargaining position, individuals with higher affective ToM were more likely to chose offers that were more favorable to the responder.

Result 4: *Individuals higher in shame proneness and affective ToM were more likely to follow the rule.*

A potential criticism of the previous results is that it is unclear if individuals that have higher propensity to feel shame or possesses greater ToM ability are behaving differently in the bargaining experiment due to norm adherence. It could be that measures of shame or ToM are correlated with factors different from norm or rule-following that are driving these results. To address this concern, this section examines whether shame, guilt, and ToM are in fact correlated with rule-following.

Regression results using ordinary least squares in Table 3.9, show that both shame and RMET are correlated with waiting time in the rule-following task. The longer subjects waited the more money they lost, but were told that the rule was to wait until the light turned green. Higher scores in shame significantly increased the amount of time subjects waited at the light. The RMET measure is positive and significant at the 10% level. This suggests that affective theory of mind may be related to rule-following, although the correlation is not strong.

3.5 Conclusion

Social norms are thought to have a strong influence on human behavior (Ostrom, 2000; Grant and Patil, 2012; Blau, 1995; Azar, 2007). Attempts to theoretically model adherence to social norms have often assumed that an individual's utility is directly influenced by norms (Bicchieri, 2006; Lopez-Perez, 2008; Kessler and Leider, 2012; Kimbrough and Vostroknutov, 2015). The propensity to follow social norms is an important part of the theoretical models, however, less is known about what individual differences may influence this propensity.

This paper tested the idea that moral emotions and the ability to recognize the emotions of others

Table 3.9: Predicting Waiting Time by Shame, Guilt, and ToM: Regression Results

	(1) Waiting Time	(2) Waiting Time	(3) Waiting Time
Shame	0.46***		
	(0.13)		
Guilt		0.13	
		(0.12)	
RMET			0.24*
			(0.14)
Female	1.86	2.72**	2.23*
	(1.17)	(1.18)	(1.18)
Intercept	19.74**	27.20***	30.36***
	(8.39)	(8.66)	(8.57)
N	195	196	198
R^2	0.135	0.079	0.097

Regressions are Ordinary Least Squares with standard errors in parentheses. Additional controls include session fixed effects, if English is the native language, and age.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

may be important determinants in explaining social norm adherence. The results show a strong correlation between shame proneness in both bargaining behavior and rule-following. Differences in guilt were not found to influence behavior. Variations in affective ToM appeared to influence behavior in bargaining and rule-following, although the correlation between ToM and rule-following was weak.

While this study is correlation, it adds to our understanding of why individuals differ in norm adherence. The propensity to feel moral emotions and the ability to understand others' emotions differs among individuals in the adult population. Social norms are highly context specific with different settings leading to different behavior (Bicchieri, 2006). This paper found that both shame and ToM were related to social norm adherence in two different settings. Future research should investigate whether shame and ToM influence norm adherence in environments not examined in this study.

Bibliography

Ahn, T., Lee, M., Ruttan, L., Walker, J., 2007. Asymmetric payoffs in simultaneous and sequential prisoner's dilemma games. Public Choice 132, 353–366.

Arli, D., Leo, C., Tjiptono, F., 2016. Investigating the impact of guilt and shame proneness on consumer ethics: a cross national study. International Journal of Consumer Studies 40, 2–13.

Azar, O. H., 2007. The social norm of tipping: a review. Journal of Applied Social Psychology 37, 380–402.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., 2001. The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functiong austism. Journal of Child Psychology and Psychiatry 42, 241–251.

Bartels, D. M., Kvaran, T., Nichols, S., 2013. Selfless giving. Cognition 129, 392–403.

Batson, D., 2011. These things called empathy: Eight related but distinct phenomenon. In: Decety, J., Ickes, W. (Eds.), The Social Neuroscience of Empathy. Cambridge: MIT Press.

Batson, D., Ahmad, N., 2001. Empathy-induced altruism in a prisoner's dilemma ii: what if the target of empathy has defected? European Journal of Social Psychology 31, 25–36.

Batson, D. C., Batson, J. G., Todd, M. R., Brummett, B. H., Shaw, L. L., Aldeguer, C. M. R., 1995. Empathy and the collective good: Caring for one of the others in a social dilemma. Journal of Personality and Social Psychology 68, 619–631.

Batson, D. C., Eklund, J. H., Chermok, V. L., Hoyt, J. L., Ortiz, B. G., 2007. An additional antecedent of empathic concern: Valuing the welfare of the person in need. Journal of Personality and Social Psychology 93, 65–74.

Batson, D. C., Moran, T., 1999. Empathy-induced altruism in a prisoner's dilemma. European Journal of Social Psychology 22, 474–482.

Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. Journal of Economic Theory 144, 1–35.

Bediou, B., Sacharin, V., Hill, C., Sander, D., Scherer, K. R., 2012. Sharing the fruit of labor: flexible applications of justice principles in an ultimatum game with joint-production. Social Justice Research 25, 25–50.

Bekkers, R., 2006. Traditional and health-related philanthropy: The role of resources and personality. Social Psychology Quarterly 69, 349–366.

Berthoz, S., Armony, J., Blair, R. J. R., Dolan, R., 2002. An fmri study of intentional and unintentional (embarrassing) violations of social norms. Brain 125, 1696–1708.

Bicchieri, C., 2006. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press.

Blanco, M., Engelmann, D., Koch, A. K., Normann, H. T., 2014. Preferences and beliefs in a sequential social dilemma: a within-subject analysis. Games and Economic Behavior 87, 122–135.

Blanco, M., Engelmann, D., Normann, H. T., 2011. A within-subject analysis of other-regarding preferences. Games and Economic Behavior 72, 321–338.

Blau, G., 1995. Influence of group lateness on individual lateness: a cross-level examination. Academy of Management Journal 38, 1483–1496.

Bolle, F., Ockenfels, P., 1990. Prisoners' dilemma as a game with incomplete information. Journal of Economic Psychology 11, 69–84.

Bolton, G. E., Brandts, J., Okenfels, A., 1998. Measuring motivations for the reciprocal responses observed in a simple dilemma game. Experimental Economics 1 (3), 207–219.

Bolton, G. E., Okenfels, A., March 2000. Erc: A theory of equity, reciprocity, and competition. The American Economic Review 90 (1), 166–193.

Bolton, G. E., Okenfels, A., 2005. A stress test of fairness measures in models of social utility. Economic Theory 25, 2005.

Bracht, J., Regner, T., 2013. Moral emotions and partnership. Journal of Economic Psychology 39, 313–326.

Bruguier, A. J., Quartz, S. R., Bossaerts, P. L., 2010. Exploring the nature of ”trading intuition”. Journal of Finance 65, 1703–1723.

Camerer, C. F., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press.

Carpenter, T. P., Tignor, S. M., Tsang, J.-A., Willett, A., 2016. Dispositional self-forgiveness, guilt- and shame- proneness, and the roles of motivational tendencies. Personality and Individual Differences 98, 53–61.

Charness, G., July 2004. Attribution and reciprocity in an experimental labor market. Journal of Labor Economics 22 (3), 665–688.

Charness, G., Dufwenberg, M., 2006. Promises and partnership. Econometrica 74, 1579–1601.

Charness, G., Levine, D. I., 2007. Intention and stochastic outcomes: An experimental study. The Economic Journal 117, 1051–1072.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. The Quarterly Journal of Economics 117, 817–869.

Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Experimental Economics 14, 47–83.

Chaudhuri, A., Paichayontvijit, T., 2006. Conditional cooperation and voluntary contributions to a public good. Economics Bulletin 3, 1–14.

Christie, R., Geis, F. L., Berger, D., 1970. Studies in machiavellianism. New York: Academic Press.

Clark, K., Sefton, M., 2001. The sequential prisoner's dilemma: Evidence on reciprocation. The Economic Journal 111, 51–68.

Cohen, T. R., Panter, S. T., Insko, C. A., 2011. Introduction the gasp scale: a new measure of guilt and shame proness. Journal of Personality and Social Psychology 100, 947–966.

Cox, J. C., Deck, C. A., 2006. Assigning intentions when actions are unobservable: The impact of trembling in the trust game. Southern Economic Journal 73, 307–314.

Cox, J. C., Friedman, D., Gjerstad, S., 2007. A tractable model of reciprocity and fairness. Games and Economic Behavior 59, 17–45.

Dana, J., Weber, R. A., Kuang, J. X., 2007. Exploting moral wiggle room: experiments demonstrating an illusory preference for fairness. Economic Theory 33, 67–80.

Davis, M. H., 1983. Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology 44, 113–126.

Dhaene, G., Bouckaert, J., 2010. Sequential reciprocity in two-player, two-stage games: An experimental analysis. Games and Economic Behavior 70, 289–303.

Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. Games and Economic Behavior 47, 268–298.

Elster, J., 1989. Social norms and economic theory. American Economic Review 3, 99–117.

Ent, M. R., Baumeister, R. F., 2015. Individual differences in guilt proness affect how people respond to moral tradeoffs beetween harm avoidance and obedience to authority. Personality and Individual Differences 74, 231–234.

Falk, A., Fehr, E., Fischbacher, U., January 2003. On the nature of fair behavior. Economic Inquiry 41 (1), 20–26.

Falk, A., Fehr, E., Fischbacher, U., 2008. Testing theories of fairness- intentions matter. Games and Economic Behavior 62, 287–303.

Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games and Economic Behavior 54, 293–315.

Falk, A., Kosfeld, M., 2006. The hidden costs of control. The American Economic Review 96, 1611–1630.

Fehr, E., Schmidt, K. M., August 1999. A theory of fairness, competition, and cooperation. The Quarterly Journal of Economics 114 (3), 817–868.

Feng, C., Luo, Y., Gu, R., Broster, L. S., Shen, X., Tian, T., Luo, Y.-J., Krueger, F., 2013. The flexible fairness: equality, earned entitlement, and self-interest. PLoS ONE 8, 1–18.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171–178.

Fischbacher, U., Gachter, S., Fehr, E., 2001. Are people conditionally cooperative? evidence from a public goods experiment. Economics Letters 71, 397–404.

Fischbacher, U., Karies-Schwarz, N., Stefani, U., 2009. Performance, productivity, and fairness: an experiment on the distribution of joint production. Working Paper.

Franco-Watkins, A. M., Edwards, B. D., Jr., R. E. A., 2013. Effort and fairness in bargaining games. Journal of Behavioral Decision Making 26, 79–90.

Frederick, S., 2005. Cognitive relection and decision making. Journal of Economic Perspectives 19, 25–42.

Friedman, O., Neary, K. R., 2008. Determining who owns what: do children infer ownership from first possession? Cognition 107, 829–849.

Gachter, S., Riedl, A., 2005. Moral property rights in bargaining with infeasible claims. Management Science 51, 249–263.

Ganter, A., Guth, W., Konigstein, M., 2001. Equitable choice in bargain games with joint production. Journal of Economic Behavior & Organization 46, 209–225.

Garcia-Gallego, A., Georgantzis, N., Jaramillo-Gutierrez, A., 2008. Ultimatum salary bargaining with real effort. Economis Letters 98, 78–83.

Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality. Games and Economic Behavior 1, 60–79.

Georganas, S., Healy, P. J., Weber, R. A., 2015. On the persistence of strategic sophistication. Journal of Economic Theory 159, 369–400.

Gill, D., Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. American Economic Review 102, 493–503.

Golan, O., Baron-Cohen, S., Hill, J., 2006. The cambridge mindreamind (cam) face-voice battery: Testing complex emotion recognition in adults with and without asperger syndrome. Journal of Autism and Developmental Disorders 36 (2), 169–83.

Grant, A. M., Patil, S. V., 2012. Challenging the norm of self-interest: minminor influence and transitions to helping norms in work units. Academy of Management Review 37, 547–568.

Guth, W., Kocher, M. G., 2014. More than thirty years of ultimatum bargaining experiments: motives, variations, and a survey of the recent literature. Journal of Economic Behavior & Organization 108, 396–409.

Hay, D. F., Ross, H. S., 1982. The social nature of early conflict. Child Development 53, 105–113.

Herrmann, B., Thoni, C., 2009. Measuring conditional cooperation: a replication study in russia. Experimental Economics 12, 87–92.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, property rights, and anonymity in bargaining games. Gams and Economic Behavior 7, 346–380.

Howell, A. J., Turowski, J. B., Buro, K., 2012. Guilt, empathy, and apology. Personality and Individual Differences 53, 917–922.

Jordan, J., Flynn, F. J., Cohen, T. R., 2015. Forgive them for i have sinned: The relationship between guilt and forgiveness of others' transgressions. European Journal of Social Psychology 45, 441–459.

Joseph Henrich, Robert Boyd, S. B. C. C. E. F. H. G., McElreath, R., 2001. In search of homo ececonomic: behavioral experiments in 15 small-scale societies. American Economic Review 91, 73–78.

Karagozoglu, E., 2012. Bargaining games with joint production. Oxford Handbook of Economic Conflict Resoltuion. Oxford University Press, New York.

Karagozoglu, E., Riedl, A., 2014. Performance information, production uncertainty, and subsubject entitlements in bargaining. Management Science Articles in Advance, 1–16.

Kessler, J. B., Leider, S., 2012. Norms and contracting. Management Science 58, 62–77.

Kimbrough, E. O., Vostroknutov, A., 2015. Norms make preferences social. Journal of the European Economic Association, 1–31.

Kirkland, R. A., Peterson, E., Baker, C. A., Miller, S., Pulos, S., 2013. Meta-analysis reveals adult female superiority in "reading the mind in the eyes test". North American Journal of Psychology 15 (1), 121–146.

Konow, J., 2000. Fair shares: Accountability and cognitive dissonance in allocation decisions. American Economic Review 90, 1072–1091.

Kruger, D. J., 2003. Evolution and altruism: Combining psychological mediators with naturally selected tendences. Evolution and Human Behavior 24, 118–125.

Leliveld, M. C., van Dijk, E., llja van Beest, 2008. Initial ownership in bargaining- introducing the giving, splitting, and taking ultimatum bargaining game. Personality and Social Psychology Bulletin 34, 1214–1225.

Leliveld, M. C., Vandijk, E., Vanbeest, I., 2012. Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive justice. European Journal of Social Psychology 42, 135–140.

Levine, D. K., 1998. Modeling altruism and spitefulness in experiments. Review of Economic Dynamics 1, 593–622.

Lewis, H. B., 1971. Shame and Guilt in Neurosis. New York: International Universities Press.

Lopez-Perez, R., 2008. Aversion to norm-breaking: a model. Games and Economic Behavior 64, 237–267.

Loveland, K. A., Pearson, D. A., Tunali-Kotoski, B., Ortegon, J., Gibbs, M. C., 2001. Judgments of social appropriateness by children and adolescents with autism. Journal of Autism and Developmental Disorders 41, 367–376.

Martino, B. D., O'Doherty, J. P., Ray, D., Bossaerts, P., Camerer, C., 2013. In the mind of the market: Theory of mind biases value computation during financial bubles. Neuron 80 (4), 1222–1231.

McCabe, K. A., Rigdon, M. L., Smith, V. L., 2003. Positive reciprocity and intentions in trust games. Journal of Economic Behavior & Organization 52, 267–275.

McCabe, K. A., Smith, V. L., LePore, M., 2000. Intentionality detection and "mindreading": Why does game form matter? Proceedings of the National Academy of Sciences of the United States of America 97, 4409–4409.

Oceja, L., Jimenez, I., 2007. Beyond egoism and gruop identity: Empathy toward the other and awareness of others in a social dilemma. The Spanish Journal of Psychology 10, 369–379.

Ostrom, E., 2000. Collective action and the evolution of social norms. Journal of Economic Perspectives 14, 137–158.

Rabin, M., 1993. Incorporating fairness into game theory and economics. The American Economic Review 83, 1281–1302.

Rand, D. G., Greene, J. D., Nowak, M. A., 2012. Spontaneous giving and calculated greed. Nature 489, 427–430.

Ridinger, G., 2011. Empathetic concern, altruism, and the pursuit of distributive justice. Master's thesis, California State University, Fullerton.

Ridinger, G., 2015a. Intentions versus outcomes: cooperation and fairness in a sequential prisoner's dilemma with nature. Working Paper.

Ridinger, G., 2015b. Ownership, punishment, and intentions in a real-effort bargaining experiment. Working Paper.

Ridinger, G., McBride, M., 2015. Money affects theory of mind differently by gender. PLOS ONE 10, e0143973.

Ross, H. S., 2013. Effects of ownership rights on cconflict between toddler peers. Infancy 18 (256-275).

Rumble, A. C., Lange, P. A. M. V., Parks, C. D., 2010. The benefits of empathy: When empathy may sustain cooperation in social dilemmas. European Journal of Social Psychology 40, 856–866.

Rustagi, D., Engel, S., Kosfeld, M., 2010. Conditional coopeation and cocost monitorying explain success in forest ccommon management. Science 330, 961–965.

Sebald, A., 2010. Attribution and reciprocity. Games and Economic Behavior 68, 339–352.

Sher, I., Koenig, M., Rustichini, A., 2014. Children's strategic theory of mind. Proceedings of the National Academy of Sciences of the United States of America 111, 13307–13312.

Singer, T., 2009. Neuroeconomics: Decision Making and the Brain. London: Academic Press, Ch. Understanding Others: Brain Mechanisms of Theory of Mind and Empathy, pp. 251–265.

Singer, T., Fehr, E., 2005. The neuroeconomics of mind reading and empathy. American Economic Review 95, 340–345.

Singer, T., Steinbeis, N., 2009. Differential roles of fairness- and compassion-based motivations for cooperation, defection, and punishment. Annals of the New York Academy of Sciences 1167, 41–50.

Smith, A., 1790. The theory of moral sentiments. London: A. Millar. Library of Economics and Liberty. Retreived September 24, 2014. `http://www.econlib.org/library/Smith/smMS.html`.

Stanca, L., 2010. How to be kind? outcomes versus intentions as determinants of fairness. Economics Letters 106, 19–21.

Tagney, J. P., Miller, R. S., Flicker, L., Barlow, D. H., 1996. Are shame, guilt, and embarrassment distinct emotions? Journal of Personality and Social Psychology 70, 1256–1269.

Tagney, J. P., Stuewig, J., Mashek, D. J., 2007. Moral emotions and moral behavior. Annual Review of Psychology 58, 345–372.

Torralva, T., Kipps, C. M., Hodges, J. R., Clark, L., Bekinschtein, T., Roca, M., maria Lujan Calcagno, Manes, F., 2007. The relationship between affaffect decision-making and theory of mind in the frontal variant of fronto-temporal dementia. Neuropsychologia 45 (2), 342–349.

Tracy, J. L., Robins, R. W., 2004. Putting the self into self-conscious emotions: a thoretical model. Psychological Inquiry 15, 103–125.

Wolf, S. T., Cohen, T. R., Panter, A., Insko, C. A., 2010. Shame proneness and guilt proneness: toward the further understanding of reactions to public and private transgressions. Self and Identity 9, 337–362.

Appendices

A Appendix

A.1 Theoretical Models of Social Preferences

A.2 Inequity Aversion

Proposition 3. *For player 2, in any pure strategy subgame perfect Nash equilibrium:*

1. *If $\beta_2 > \frac{1}{3}$, then player 2 will cooperate if nature cooperates and defect if nature defects.*
2. *If $\beta_2 < \frac{1}{3}$, then player 2 will always defect.*

Proof. To see that 1 holds, first let us look at when player 1 and nature cooperates. Player 2 will cooperate if $3 > 4 - 3 \cdot \beta_2$. This occurs when $\beta_2 > \frac{1}{3}$. Due to symmetry, this condition also ensures that if player 1 defects and nature cooperates, then player 2 will cooperate. To see that player 2 will defect if nature defects, let us look at the case when player 1 cooperates and nature defects. Player 2 will choose to defect if $2 > 1 - 3 \cdot \alpha_2$ which holds for all α_2 since $\alpha_2 \in [0, 1]$. Due to symmetry we can see that this will also hold if player 1 defects and nature defects.

□

A.3 Reciprocity Model

In this section, I introduce the intention-based reciprocity model of Dufwenberg and Kirchsteiger (2004). The model uses psychological game theory based on Battigalli and Dufwenberg (2009). Psychological games, first developed by Geanakoplos et al. (1989), differ from standard games in that an individual's beliefs directly affects her utility. Using psychological game theory, Rabin (1993) modeled concerns for reciprocity in normal form games and Dufwenberg and Kirchsteiger (2004) extended the idea to extensive form games. Reciprocity is captured by the idea that people like to be kind to people who are kind to them and be unkind to people who are unkind to them.

Formally, let $I \in \{0,1,2\}$ be the set of players. Denote nature as player 0. Let H be the set of histories that lead to subgames. Each player $i \in I \setminus \{0\}$ has a set of possible strategies A_i. The strategy set is $A = \prod_{i \in I \setminus \{0\}} A_i$. Each strategy $a_i \in A_i$ gives a probability distribution on the possible choices of player i at each history $h \in H$. Each player $i's$ updated strategy is defined as $a_i(h)$.[1] The probability distribution for the behavioral strategy of the chance player is defined as θ which is commonly known to both players. Given end node payoffs, the expected material payoff for each player $i \in I \setminus \{0\}$ is $\pi_i : A \times \{\theta\} \to \Re$.

Following Dufwenberg and Kirchsteiger (2004) and Sebald (2010), additional notation must be introduced as it is necessary to keep track of first and second order beliefs. Each player i has a set of beliefs, B_{ij}, about the strategy of player j. Let $b_{ij} \in B_{ij}$ be the belief player i has about the strategy of player j. Let C_{iji} define the set of beliefs player i has about the belief player j has about player $i's$ strategy. Define $c_{iji} \in C_{iji}$ as the belief player i has about the belief player j has about player $i's$ strategy. To capture the main features of reciprocity beliefs need to be updated as the game progresses. Let $b_{ij}(h)$ and $c_{iji}(h)$ represent the updated beliefs at history h.

To capture a concern for fair outcomes this paper uses inequity aversion developed by Fehr and Schmidt (1999). Inequity averse agents care about both their own payoff and another player's

[1] Here the only difference between a_i and $a_i(h)$ is that choices in history h are made with probability one.

payoff. This can capture the idea that people might prefer outcomes that are more equal.

Combining both reciprocity and inequity aversion into a single model gives:

$$U_i(a_i(h), b_{ij}(h), c_{iji}(h), \theta) = \pi_i(a_i(h), a_j(h), \theta) \tag{A.1}$$
$$+\lambda_i \cdot k_{ij}(a_i(h), b_{ij}(h), \theta) \cdot \phi_{iji}(b_{ij}(h), c_{iji}(h), \theta)]$$

where $\lambda_i \geq 0$.

In A.1, $i's$ utility depends on $i's$ own payoff plus concerns for reciprocity . The weight that i places on reciprocity concerns is captured by λ_i. The function $k_{ij}(a_i(h), b_{ij}(h), \theta)$ is a measure of the kindness of i towards j at history h, and $\phi_{iji}(b_{ij}(h), c_{iji}(h), \theta)$ is $i's$ belief about the kindness of j towards i at history h. The kindness of player i towards player j is represented as a function of player $i's$ strategy choice $a_i(h)$ and belief $b_{ij}(h)$. At a specific $a_i(h)$ and $b_{ij}(h)$ the kindness of player i is captured by the payoff that player j gets minus an equitable payoff. The kindness function is defined as:

$$k_{ij}(a_i(h), b_{ij}(h), \theta) = \pi_j(a_i(h), b_{ij}(h), \theta) - \pi_j^{ei}((b_{ij}(h), \theta) \tag{A.2}$$

Dufwenberg and Kirchsteiger (2004) calculate the equitable payoff for player j by finding the a_i that gives player j the highest possible payoff and finding the a_i that gives player j the lowest possible payoff. The equitable payoff is an average of the payoffs for player j evaluated at each a_i. This equitable payoff is:

$$\pi_j^{ei}(b_{ij}, \theta) = \frac{1}{2}[\max_{a_i \in A_i}\{\pi_j(a_i, b_{ij}, \theta)\} + \min_{a_i \in E_i}\{\pi_j(a_i, b_{ij}, \theta)\}] \tag{A.3}$$

where E_i, defined by Dufwenberg and Kirchsteiger (2004), is the set of efficient strategies for player i such that

$$E_i = \{a_i \in A_i | \text{ there exists no } a_i' \in A_i \text{ such that for all } h \in H,$$

$$(a_j)_{j \neq i} \in \prod_{j \neq i} A_j, \text{ and } k \in I \text{ it holds that } \pi_k(a_i'(h), (a_j(h))_{j \neq i}) \geq \pi_k(a_i(h), (a_j(h))_{j \neq i}), \quad \text{(A.4)}$$

$$\text{with strict inequality for some } (h, (a_j)_{j \neq i}, k)\}.$$

Player $i's$ belief about the kindness of player j towards player i has a similar structure and is defined as[2]:

$$\phi_{iji}(b_{ij}(h), c_{iji}(h), \theta) = \pi_i(b_{ij}(h), c_{iji}(h), \varepsilon) - \pi_j^{ej}(c_{iji}(h), \theta) \quad \text{(A.5)}$$

The equilibrium concept used in this paper is the sequential reciprocity equilibrium (Dufwenberg and Kirchsteiger, 2004; Sebald, 2010). Define for all $a = (a_i)_{i \in I} \in A$ and history $h \in H$, let $A_i(a,h) \subseteq A_i$ be the set of behavioral strategies for each player i that give the same choices as the strategy $a_i(h)$ for all histories other than h.

Definition 1. The profile $a^\star = (a_i^\star)_{i \in I \setminus \{0\}}$ is a sequential reciprocity equilibrium (SRE) if for all $i \in I \setminus \{0\}$ and for each history $h \in H$ the following properties hold:

(i) $a_i^\star(h) \in \underset{a_i \in A_i(h,a)}{\operatorname{argmax}} U_i(a_i(h), b_{ij}(h), c_{iji}(h), \theta)$ where $i \neq j$

(ii) $b_{ij} = a_j^\star$ for all $j \neq i$

(iii) $c_{iji} = a_i^\star$ for all $j \neq i$

Property (i) means that at history h, player i chooses a strategy profile that maximizes $i's$ utility given $i's$ belief. In addition, it assures that player i follows the equilibrium strategy at all other histories. At the initial history, properties (ii) and (iii) imply that initial beliefs are correct. Property (i) adds that any sequence of choices that lead to a history have probability one. As a result, the SRE concept requires that in equilibrium beliefs be correct.

[2]Here the equitable payoff is mathematically equivalent to (3).

Proposition 4. *Under perfect information and if $\theta < \frac{1}{2}$, then in any SRE the potential behavior for player 2 can be described as follows:*

(a) *If $\theta < \frac{1}{4}$, and $\lambda_2 > \frac{1}{1-4\theta}$, then player 2 will cooperate if player 1 cooperates and defect if player 1 defects.*

(b) *$\lambda_2 < \frac{1}{2-4\theta}$, then player 2 will always defect.*

Proof. Player 2 can choose to cooperate or defect at each node h^3, h^4, h^5, and h^6 labeled in figure 1.1. Let player 1's belief about what player 2 will choose at each node be defined as: $x_1 = P_1(2\, choses\, C|h^3)$, $x_2 = P_1(2\, choses\, C|h^4)$, $x_3 = P_1(2\, choses\, C|h^5)$, and $x_4 = P_1(2\, choses\, C|h^6)$. Player 2's belief about player 1's belief about what player 2 will choose at each node is defined as the expectation of player 1's beliefs about player 2. This gives: $y_1 = E_2[x_1|h^3]$, $y_2 = E_2[x_2|h^4]$, $y_3 = E_2[x_3|h^5]$, and $y_4 = E_2[x_4|h^6]$. Player 1 can choose to cooperate or defect at node h^0. Let player 2's belief that player 1 will cooperate be $z_1 = P_2(1\, choses\, C|h^0)$. Player 1's belief about player 2's belief that player 1 will cooperate is defined as $w_1 = E_1[z_1|h^0]$. The game can now be analyzed as a psychological game with reciprocity.

If player 1 cooperates, then player 2's belief about the the kindness of player 1 towards player 2 is $\phi_{212} = \frac{1}{2}((1-\theta)(4-y_1)+\theta(2-y_2)-\theta(4-y_3)-(1-\theta)(2-y_4))$. If player 1 defects, then $\phi_{212} = \frac{1}{2}(\theta(4-y_3)+(1-\theta)(2-y_4)-(1-\theta)(4-y_1)-(\theta)(2-y_2))$. In any SRE player 2 will always make the same decision at history h^5 and h^6. To see why, note that for player 2 to defect at h^6 it must be that $1+\lambda_2[(1-\theta)(4-y_1)+\theta(2-y_2)-\theta(4-y_3)-(1-\theta)(2-y_4)] > 0$. In order for the second mover to defect at h^5 it must be that $1+\lambda_2[(1-\theta)(4-y_1)+\theta(2-y_2)-\theta(4-y_3)-(1-\theta)(2-y_4)] > 0$. As a result, in any SRE it must be the case that $x_3 = x_4 = y_3 = y_4$. Similarly, in any SRE player 2 will always make the same decision at history h^1 and h^2.

If (a) holds in equilibrium, then $x_1 = x_2 = y_1 = y_2 = 1$ and $x_3 = x_4 = y_3 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(1-4\theta)$. At h^3 player 2 will cooperate if $3+(\frac{1}{2})\lambda_2(1-\gamma_2)(1-4\theta) >$

$4-(\frac{1}{2})\lambda_2(1-\gamma_2)(1-4\theta)$. This holds if $\theta < \frac{1}{4}$ and $\lambda_2 > \frac{1}{1-4\theta}$. At h^4 , player 2 will cooperate if $\lambda_2(1-4\theta) > 1$. This holds if $\theta < \frac{1}{4}$ and $\lambda_2 > \frac{1}{1-4\theta}$. Since $\gamma_2 \in [0,1]$, then in order for player 2 to cooperate in pure strategies at h^4 it must be the case that $\theta > \frac{1}{4}$. For player 2 to defect at h^5, then the following must hold $\lambda_2(1-4\theta) > -1$. Since $\lambda_2(1-4\theta) > 1$, then player 2 will defect at h^5. As a result, if $\theta > \frac{1}{4}$, and $\lambda_2 > \frac{1}{1-4\theta}$, then player 2 will cooperate if player 1 cooperates and defect if player 1 defects.

For (b) it must be the case that $y_1 = y_2 = y_3 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(2-4\theta)$. At h^3 player 2 will defect if $\lambda_2(2-4\theta)+ < 1$. This holds if $\lambda_2 < \frac{1}{2-4\theta}$. At h^4 player 2 will defect if $\lambda_2(2-4\theta) < 1$ which holds if $\lambda_2 < \frac{1}{2-4\theta}$. At h^5 player 2 will defect if $\lambda_2(2-4\theta) > -1$. As a result, if $\lambda < \frac{1}{2-4\theta}$, then player 2 will always defect. □

Proposition 5. *Under imperfect information and if $\theta < \frac{1}{2}$, then in any SRE the potential behavior for player 2 can be described as follows:*

1. *If player 1 cooperates*
   (a) $\lambda_2 > \frac{1}{2-4\theta}$ *, then player 2 will always cooperate.*
   (b) $\lambda_2 < \frac{1}{2-4\theta}$*, then player 2 will always defect.*
2. *If player 1 defects, then player 2 will always defect.*

Proof. Let player 1's belief about what player 2 will choose at each information set be defined as: $q_1 = P_1(2\,choses\,C|h^3 \cup h^5)$, and $q_2 = P_1(2\,choses\,C|h^4 \cup h^6)$. Player 2's belief about player 1's belief about what player 2 will choose at each information set is defined as the expectation of player 1's beliefs about player 2. This gives: $v_1 = E_2[q_1|h^3 \cup h^5]$, and $v_2 = E_2[q_2|h^4 \cup h^6]$. Player 1 can choose to cooperate or defect at node h^0. Let player 2's belief that player 1 will cooperate be $z_1 = P_2(1\,choses\,C|h^0)$. Player 1's belief about player 2's belief that player 1 will cooperate is defined as $w_1 = E_1[z_1|h^0]$.

Player 2 only observes the results of nature. Player 2's evaluation of the kindness of player 1 depends on the belief about what node she is currently at. If player 2 observes cooperation, then the probability that player 2 believes she is at node h^3 is $P(h^3|2\,observes\,C) = \frac{z_1 \cdot \varepsilon_1}{z_1 \cdot \varepsilon_1 + (1-z_1) \cdot \varepsilon_2}$ via Bayes rule. Similarly, $P(h^5|2\,observes\,C) = \frac{(1-z_1) \cdot \varepsilon_2}{z_1 \cdot \varepsilon_1 + (1-z_1) \cdot \varepsilon_2}$. If player 2 observes defection, then $P(h^4|2\,observes\,D) = \frac{z_1 \cdot (1-\varepsilon_1)}{z_1 \cdot (1-\varepsilon_1) + (1-z_1) \cdot (1-\varepsilon_2)}$ and $P(h^6|2\,observes\,D) = \frac{(1-z_1) \cdot (1-\varepsilon_2)}{z_1 \cdot (1-\varepsilon_1) + (1-z_1) \cdot (1-\varepsilon_2)}$. Since the SRE concept requires that initial beliefs be correct, it follows that player 2 knows in equilibrium what player 1 chooses.

Player 2's belief about the kindness of player 1 is $\phi_{212} = (z_1 - \frac{1}{2})(1-2\theta)(2 - v_1 + v_2)$. No matter the history, the kindness of player 2 towards player 1 will always be $k_{21} = 1$ if player 2 cooperates and $k_{21} = -1$ if player 2 defects.

Suppose that in equilibrium player 1 cooperates, then player 2 knows this. Since $z_1 = w_1 = 1$, player 2 knows that if nature cooperates then she is at node $h^3$and if nature defects then she is at node h^4. If nature cooperates, then player 2 will cooperate if $\lambda_2(1-2\theta)(2-v_1+v_2) > 1$. If nature defects, then player 2 will cooperate if $\lambda_2(1-2\theta)(2-v_1+v_2) > 1$. As a result, player 2 will make the same decision at nodes h^3 and h^4. Similarly, if in equilibrium player 1 defects, $z_1 = w_1 = 0$. In this case, player 2 will make the same decision at nodes h^4 and h^5.

For 1(a), $v_1 = v_2 = 1$. This is possible if $\lambda_2(1-2\theta)(2) > 1$. If $\lambda_2 > \frac{1}{2-4\theta}$, then player 1 will cooperate no matter the results of nature.

For 1(c), $v_1 = v_2 = 0$. This is possible if both $\lambda_2(1-2\theta)(2) < 1$. If $\lambda_2 < \frac{1}{2-4\theta}$, then player 2 will always defect.

For 2, $v_1 = v_2 = 0$. This is possible if $-\lambda_2(1-2\theta)(2) < 1$, which holds for all $\lambda_2 \geq 0$. As a result, if player 1 defects, then player 2 will defect. □

A.4 Mixed-concerns Model

In this section, I introduce the mixed-concerns model that combines the models of inequity aversion Fehr and Schmidt (1999), and reciprocity Dufwenberg and Kirchsteiger (2004) into a single framework. Let the utility of an individual i be defined as:

$$\begin{aligned} U_i(a_i(h), b_{ij}(h), c_{iji}(h), \theta) &= \pi_i(a_i(h), a_j(h), \theta) \\ &+\rho_i \cdot [(1-\gamma_i) \cdot k_{ij}(a_i(h), b_{ij}(h), \theta) \cdot \phi_{iji}(b_{ij}(h), c_{iji}(h), \theta) + \gamma_i \cdot D_{ij}(a_i(h), b_{ij}(h), \theta)] \end{aligned} \tag{A.6}$$

where $\rho_i \geq 0$ and $\lambda_i \in [0,1]$. In A.6, $i's$ utility depends on $i's$ own payoff plus concerns for reciprocity and inequity aversion. The weight that i places on these social preferences is captured by ρ_i. An additional parameter, γ_i, is the relative weight placed on concerns for reciprocity and distribution. Higher values of γ_i mean that person i places a lower weight on reciprocity and greater weight on distributional concerns.

The function $D_{ij}(a_i(h), b_{ij}(h), \theta)$ captures the distributional concerns of an individual. $D_{ij}(a_i(h), b_{ij}(h), \theta)$ is assumed to be a modified version of inequity aversion defined as:

$$D_{ij}(a_i(h), b_{ij}(h), \varepsilon) = -\max\{\pi_j - \pi_i,\ \pi_i - \pi_j\}$$

where $\pi_j = \pi_j(a_i(h), b_{ij}(h), \theta)$ and $\pi_i = \pi_i(a_i(h), b_{ij}(h), \theta)$.[3] The functional form for $D_{ij}(a_i(h), b_{ij}(h),)$ does not capture the idea from the inequity aversion model that people might dislike getting less than another person more than they feel bad about getting more. This could easily be incorporated into the model, but has been left out for simplicity.[4]

The function $k_{ij}(a_i(h), b_{ij}(h), \theta)$ is a measure of the kindness of i towards j at history h, and

[3] Many different types of distributional concerns could be considered. Other forms to be included could be Rawlsian, Utilitarian, or Nash Product.

[4] Assuming the standard function form for inequity aversion Fehr and Schmidt (1999) gives the same equilibrium predictions for second movers in the sequential prisoner's dilemma with nature as the restricted functional form assumed here.

$\phi_{iji}(b_{ij}(h), c_{iji}(h), \theta)$ is $i's$ belief about the kindness of j towards i at history h. Both $k_{ij}(a_i(h), b_{ij}(h), \theta)$ and $\phi_{iji}(b_{ij}(h), c_{iji}(h), \theta)$ have the same functional form described in the previous section. Since the focus is on sequential games, the analysis uses the sequential reciprocity equilibrium as it allows beliefs to be updated.

One advantage of the mixed-concerns model compared to other models that combine concerns for intentions and outcomes (Charness and Rabin, 2002; Falk and Fischbacher, 2006), is that chance players are incorporated into the model. Chance players are often used in theoretical models to capture many different environments and random devices are often used in experiments. The mixed-concerns model can make equilibrium predictions in these situations. In addition, the model allows us to investigate how changes in the distribution of the choices by chance players influences equilibrium predictions.

Perfect Information

The main focus on this analysis will be on what player 2 chooses to do in the game with perfect information. In any SRE, the potential behavior for player 2 is described in proposition 1. Due to player 2's behavior being the primary focus of the paper, player 1's equilibrium behavior is left to the Supplemental Appendix.

Proposition 6. *If $\theta < \frac{1}{2}$, then in any SRE the potential behavior for player 2 can be described as follows:*

(a) *If $\theta < \frac{1}{4}$, $0 \leq \gamma_2 \leq \frac{1-4\theta}{4-4\theta}$, and $\rho_2 > \frac{1}{1-4\theta-\gamma_2(4-4\theta)}$, then player 2 will cooperate if player 1 cooperates and defect if player 1 defects.*

(b) *If $\theta < \frac{1}{3}$, $0 < \gamma_2 < \frac{1-3\theta}{4-3\theta}$, and $\frac{1}{1-3\theta+\gamma_2(2+2\theta)} < \rho_2 < \frac{1}{1-3\theta-\gamma_2(2-4\theta)}$, then player 2 will cooperate if player 1 and nature cooperates, and defect otherwise.*

(c) *If* $\frac{1-2\theta}{4-2\theta} < \gamma_2$ *and* $\rho_2 > \frac{1}{\gamma_2(2+2\theta)-1+2\theta}$, *then player 2 will cooperate if nature cooperates and defect if nature defects.*

(d) *If* $\gamma_2 > \frac{2-4\theta}{5-4\theta}$ *and* $\rho_2 < \frac{1}{2-4\theta+\gamma_2(4\theta-1)}$, *or* $\gamma_2 < \frac{2-4\theta}{5-4\theta}$ *and* $\rho_2 < \frac{1}{2-4\theta+\gamma_2(5-4\theta)}$, *then player 2 will always defect.*

Proof. Player 2 can choose to cooperate or defect at each node h^3, h^4, h^5, and h^6 labeled in figure 1.1. Let player 1's belief about what player 2 will choose at each node be defined as: $x_1 = P_1(2\, choses\, C|h^3)$, $x_2 = P_1(2\, choses\, C|h^4)$, $x_3 = P_1(2\, choses\, C|h^5)$, and $x_4 = P_1(2\, choses\, C|h^6)$. Player 2's belief about player 1's belief about what player 2 will choose at each node is defined as the expectation of player 1's beliefs about player 2. This gives: $y_1 = E_2[x_1|h^3]$, $y_2 = E_2[x_2|h^4]$, $y_3 = E_2[x_3|h^5]$, and $y_4 = E_2[x_4|h^6]$. Player 1 can choose to cooperate or defect at node h^0. Let player 2's belief that player 1 will cooperate be $z_1 = P_2(1\, choses\, C|h^0)$. Player 1's belief about player 2's belief that player 1 will cooperate is defined as $w_1 = E_1[z_1|h^0]$. The game can now be analyzed as a psychological game with mixed concerns.

If player 1 cooperates, then player 2's belief about the the kindness of player 1 towards player 2 is $\phi_{212} = \frac{1}{2}((1-\theta)(4-y_1)+\theta(2-y_2)-\theta(4-y_3)-(1-\theta)(2-y_4))$. If player 1 defects, then $\phi_{212} = \frac{1}{2}(\theta(4-y_3)+(1-\theta)(2-y_4)-(1-\theta)(4-y_1)-(\theta)(2-y_2))$. In any SRE player 2 will always defect at history h^6. To see why, note that for player 2 to defect at h^6 it must be that $1+\rho_2(1-\gamma_2)[(1-\theta)(4-y_1)+\theta(2-y_2)-\theta(4-y_3)-(1-\theta)(2-y_4)]+3\rho_2\gamma_2 > 0$. This holds if $\rho_2 \geq 0$ and $\gamma_2 \geq 0$. As a result, in any SRE it must be the case that $y_4 = x_4 = 0$. In addition, player 2 will not cooperate at both h^4 and h^5. In order for cooperate to hold at both of those nodes, it would have to be that $\rho_2(1-\gamma_2)\phi_{212} > 1+\rho_2\gamma_2\cdot 3$ and $\rho_2\gamma_2\cdot 3 > 1+\rho_2(1-\gamma_2)\phi_{212}$ which cannot occur. As a result, an equilibrium where player 2 cooperates at both h^4 and h^5 can be ruled out.

If (a) holds in equilibrium, then $x_1 = x_2 = y_1 = y_2 = 1$ and $x_3 = x_4 = y_3 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(2(1-2\theta)-1)$. At h^3 player 2 will cooperate if $3+(\frac{1}{2})\rho_2(1-\gamma_2)(2(1-2\theta)-1) > 4-(\frac{1}{2})\rho_2(1-\gamma_2)(2(1-2\theta)-1)-\rho_2\cdot\gamma_2\cdot 3$, where $D_{21} = 0$ if player 2 cooperates and

$D_{21} = 3$ if player 2 defects. This holds if $\gamma_2 \geq \frac{1-2(1-2\theta))}{4-2(1-2\theta)}$ and $\rho_2 > \frac{1}{2(1-2\theta)-1+\gamma_2(4-2(1-2\theta))}$. At h^4 , player 2 will cooperate if $\rho_2(1-\gamma_2)(2(1-2\theta)-1)-\rho_2 \cdot \gamma_2 \cdot 3 > 1$. This holds if $\gamma_2 \leq \frac{2(1-2\theta)-1}{2(1-2\theta)+2}$, and $\rho_2 > \frac{1}{2(1-2\theta)-1-\gamma_2(2(1-2\theta)+2)}$. Since $\gamma_2 \in [0,1]$, then in order for player 2 to cooperate in pure strategies at h^4 it must be the case that $\theta > \frac{1}{4}$. Since player 2 cooperated at h^4, then it must be the case that player 2 defects at h^5. For player 2 to defect at h^5, then the following must hold $\rho_2(1-\gamma_2)(2(1-2\theta)-1)-\rho_2 \cdot \gamma_2 \cdot 3 > -1$. Since $\rho_2(1-\gamma_2)(2(1-2\theta)-1)-\rho_2 \cdot \gamma_2 \cdot 3 > 1$, then player 2 will defect at h^5. As a result, if $\theta > \frac{1}{4}$, $0 \leq \gamma_2 \leq \frac{1-4\theta}{4-4\theta}$, and $\rho_2 > \frac{1}{1-4\theta-\gamma_2(4-4\theta)}$, then player 2 will cooperate if player 1 cooperates and defect if player 1 defects.

For (b), in equilibrium it must be the case that $y_1 = 1$ and $y_2 = y_3 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(1-3\theta)$. At h^3 player 2 will cooperate if $3+\rho_2(1-\gamma_2)(1-3\theta) > 4-\rho_2 \cdot \gamma_2 \cdot 3$. This holds if $\gamma_2 > \frac{3\theta-1}{2+3\theta}$ and $\rho_2 > \frac{1}{1-3\theta+\gamma_2(2+3\theta))}$. At h^4 player 2 will defect if $1 > \rho_2(1-\gamma_2)(1-3\theta) - \rho_2 \cdot \gamma_2 \cdot 3$. This holds if $\gamma_2 < \frac{1-3\theta}{4-3\theta}$ and $\rho_2 < \frac{1}{1-3\theta-\gamma_2(4-3\theta)}$. Since $\gamma_2 \in [0,1]$, in order for player 2 to defect at h^4, then it must be the case that $\theta < \frac{1}{3}$. In order for player 2 to defect at h^5, then it must be the case that $\rho_2(1-\gamma_2)(1-3\theta)-\rho_2 \cdot \gamma_2 \cdot 3 > -1$. This holds if $\rho_2 > \frac{-1}{1-3\theta-\gamma_2(4-3\theta)}$ and $\gamma_2 \leq \frac{1-3\theta}{4-3\theta}$. As a result, if $\theta < \frac{1}{3}$, $0 < \gamma_2 < \frac{1-3\theta}{4-3\theta}$, $\frac{1}{1-3\theta+\gamma_2(2+2\theta)} < \rho_2 < \frac{1}{1-3\theta-\gamma_2(2-4\theta)}$, then player 2 will cooperate if player 1 and nature cooperates and defect otherwise.

For (c), in equilibrium beliefs must be correct, which gives $y_1 = y_3 = 1$ and $y_2 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(1-2\theta)$. At h^3 player 2 will cooperate if $\rho_2(1-\gamma_2)(1-2\theta) + \rho_2 \cdot \gamma_2 \cdot 3 > 1$. This holds if $\gamma_2 > \frac{2\theta-1}{2+2\theta}$ and $\rho_2 > \frac{1}{1-2\theta+\gamma_2(2+2\theta)}$. Player 2 will defect at h^4 if $-\rho_2(1-\gamma_2)(1-2\theta)+\rho_2 \cdot \gamma_2 \cdot 3 > -1$. This holds if $\gamma_2 > \frac{1-2\theta}{4-2\theta}$ and $\rho_2 > \frac{-1}{2\theta-1+\gamma_2(4-2\theta)}$. Player 2 will cooperate at h^5 if $-\rho_2(1-\gamma_2)(1-2\theta)+\rho_2 \cdot \gamma_2 \cdot 3 > 1$. This will hold if $\gamma_2 > \frac{1-2\theta}{4-2\theta}$ and $\rho_2 > \frac{1}{2\theta-1+\gamma_2(4-2\theta)}$. Since $\frac{1}{2\theta-1+\gamma_2(4-2\theta)} \geq \frac{1}{1-2\theta+\gamma_2(2+2\theta)}$, and $\gamma_2 \in [0,1]$, then in order to have this equilibrium it must be the case that$\rho_2 > \frac{1}{2\theta-1+\gamma_2(4+2\theta)}$. So if $\gamma_2 > \frac{1-2\theta}{4-2\theta}$, and t$\rho_2 > \frac{1}{2\theta-1+\gamma_2(4+2\theta)}$, then player 2 will cooperate if nature cooperates and defect if nature defects.

For (d) it must be the case that $y_1 = y_2 = y_3 = y_4 = 0$. If player 1 cooperates, then $\phi_{212} = \frac{1}{2}(2-4\theta)$. At h^3 player 2 will defect if $\rho_2(1-\gamma_2)(2-4\theta)+\rho_2 \cdot \gamma_2 \cdot 3 < 1$. This holds if $\gamma_2 \geq 0$ and $\rho_2 <$

$\frac{1}{2-4\theta+\gamma_2(4\theta-1)}$. At h^4 player 2 will defect if $\rho_2(1-\gamma_2)(2-4\theta))-\rho_2\cdot\gamma_2\cdot 3<1$ which holds if $\gamma_2<\frac{2-4\theta}{5-4\theta}$ and $\rho_2<\frac{1}{2-4\theta-\gamma_2(5-4\theta)}$ or if $\gamma_2\geq\frac{2-4\theta}{5-4\theta}$ and $\rho_2\geq 0$. At h^5player 2 will defect if $\rho_2(1-\gamma_2)(2-4\theta)-\rho_2\cdot\gamma_2\cdot 3>-1$. This holds if $\gamma_2>\frac{2-4\theta}{5-4\theta}$ and $\rho_2<\frac{1}{4\theta-2+\gamma_2(5-4\theta)}$ or if$\gamma_2<\frac{2-4\theta}{5-4\theta}$ and $\rho_2\geq 0$. If $\gamma_2<\frac{2-4\theta}{5-4\theta}$, then $\frac{1}{2-4\theta-\gamma_2(5-4\theta)}>\frac{1}{2-4\theta+\gamma_2(5-4\theta)}$. If$\gamma_2>\frac{2-4\theta}{5-4\theta}$, then $\frac{1}{4\theta-2+\gamma_2(5-4\theta)}>\frac{1}{2-4\theta+\gamma_2(4\theta-1)}$. Given this, it follows that if $\gamma_2>\frac{2-4\theta}{5-4\theta}$, and $\rho_2<\frac{1}{2-4\theta+\gamma_2(4\theta-1)}$, then player 2 will always defect. If $\gamma_2<\frac{2-4\theta}{5-4\theta}$ and $\rho_2<\frac{1}{2-4\theta+\gamma_2(5-4\theta)}$, then player 2 will always defect. □

Concerns for only reciprocity (Dufwenberg and Kirchsteiger, 2004) or only inequity aversion (Fehr and Schmidt, 1999) arise as special cases. In order to understand the differences between the models of reciprocity and inequity aversion assume that $\gamma_i=0$. In other words, assume that players are purely reciprocal. As a result of proposition 1, if player 1 cooperates, then player 2 will cooperate if $\rho_2>\frac{1}{1-4\theta}$ and $\theta<\frac{1}{4}$. This implies that conditional cooperation by player 2 is only possible provided that player 1's control is sufficiently high. If $\theta>\frac{1}{4}$, then player 2 will not cooperate in pure strategies if player 1 cooperates. For player 2 to interpret a choice by player 1 as kind or unkind, player 1 has to have a certain amount of control over that choice. This model suggests that when control is low, reciprocity is not sufficient to maintain cooperation by player 2. Note that as θ decreases, then ρ_2 must be lower in order to sustain defection as a pure strategy SRE. In other words, as control by player 1 increases, lower concerns for reciprocity are needed for player 2 to always defect.

If players are only inequity averse, then this implies that $\gamma_i=1$. As result of proposition 1, inequity aversion predicts that player 2's choice is not influenced by the values of θ. The intended choice of player 1 does not influence what player 2 will choose. Player 2's choice depends only on the degree to which player 2 dislikes getting more than player 1. Cooperation by player 2 is determined by whether player 2 feels “guilty” over receiving more than player 1. If $\rho_2>\frac{1}{3}$, then player 2 will cooperate if nature cooperates regardless of player 1's choice. That is, the intended choice by player 1 is not behaviorally relevant. This contrasts with the pure reciprocity case in which player 1's intended choice matters for player 2 rather than the results of nature.

If players instead have mixed concerns about reciprocity and inequity aversion, then there are four possible pure strategy equilibria that could hold. If the equilibrium (a) occurs, then player 2 is more concerned about player 1's intentions. This leads to reciprocal behavior where player 2 cooperates if player 1 cooperates and defects if player 1 defects. Provided player 1 has a sufficient level of control over the outcome, this equilibrium is possible. Notice that this equilibrium depends upon player 2's concern for inequity aversion. Lower values of γ_2 suggest that player 2 is more reciprocal; however, if γ_2 is large, then this equilibrium may not occur due to the strong preference for equal outcomes.

The mixed concerns model also suggests another possible equilibrium (b). Here player 2 cooperates if player 1 and nature cooperates, but defects at all other histories. This equilibrium is not possible in the cases of pure reciprocity or pure inequity aversion. In this equilibrium, player 2 cooperates only if player 1 intended to cooperate and the result of that intention leads to cooperation. Intentions are not enough for player 2 to cooperate when player 1 cooperates and nature defects. In addition, if the concern about inequity aversion is sufficiently small, then player 2 will not cooperate if player 1 defects and nature cooperates. Here player 2 may be concerned about both intentions and the distribution of outcomes, but cooperation is only sustained when those concerns align.

The equilibrium (c) occurs if players are strongly inequity aversion averse. One thing to notice is that this equilibrium has no restrictions on the value of θ other than the assumption that $\theta > \frac{1}{2}$. Since the mixed concerns model allows inequity aversion and reciprocity, player 2 must have a sufficiently high concern for inequity aversion in order for (3) to hold. One interesting result is that as the value of θ increases, this equilibrium holds for smaller values of γ_2. This result makes intuitive sense. To see why, suppose that player 2 is really concerned about reciprocity. When player 1 has little control, player 1's choice is not seen as very intentional. Consequently, reciprocity has little weight in player 2's decision. As a result, inequity aversion can become more important as first mover control decreases. Since reciprocity is not much of a factor when control is low, concerns for reciprocity do not conflict as much with concerns for inequity aversion at nodes h^4 and h^5.

The equilibrium (d) gives the case when player 2 will always defect. If $\rho_2 = 0$, then the model is just the self-interest model and player 2 will always defect. If $\rho_2 > 0$, then the minimum value of ρ_2 that will lead to player 2 always defecting depends on the relative weight they place on the two concerns and the reversal probability.

Imperfect Information

In the imperfect information game, player 2 does not know what player 1 chose but does know the results of nature. In any SRE, the potential behavior for player 2 is described in proposition 2.

Proposition 7. *If $\theta < \frac{1}{2}$, then in any SRE the potential behavior for player 2 can be described as follows:*

1. *If player 1 cooperates*
   (a) *If $0 < \gamma_2 < \frac{1-2\theta}{4-2\theta}$ and $\frac{1}{1-2\theta+\gamma_2(2+2\theta)} < \rho_2 < \frac{1}{1-2\theta-\gamma_2(4-2\theta)}$ or $\gamma_2 > \frac{1-2\theta}{4-2\theta}$ and $\rho_2 > \frac{1}{1-2\theta+\gamma_2(2+2\theta)}$, then player 2 will cooperate if nature cooperates and defect if nature defects.*
   (b) *If $\gamma_2 < \frac{2-4\theta}{5-4\theta}$ and $\rho_2 > \frac{1}{2-4\theta-\gamma_2(5-4\theta)}$, then player 2 will always cooperate.*
   (c) If $\gamma_2 \geq 0$ and $\rho_2 < \frac{1}{2-4\theta+\gamma(1+4\theta)}$, *then player 2 will always defect.*
2. *If player 1 defects*
   (a) If $\gamma_2 > \frac{1-2\theta}{4-2\theta}$ *and* $\rho_2 > \frac{1}{\gamma_2(4-2\theta)-1+2\theta}$, *then player 2 will cooperate if nature cooperates and defect if nature defects.*
   (b) If $\gamma_2 > \frac{2-4\theta}{5-4\theta}$ *and* $\rho_2 < \frac{1}{\gamma_2(5-4\theta)-2+4\theta}$ *or* $\gamma_2 < \frac{2-4\theta}{5-4\theta}$ *and* $\rho_2 \geq 0$, *then player 2 will always defect.*

Proof. Let player 1's belief about what player 2 will choose at each information set be defined as: $q_1 = P_1(2\, choses\, C|h^3 \cup h^5)$, and $q_2 = P_1(2\, choses\, C|h^4 \cup h^6)$. Player 2's belief about player 1's belief about what player 2 will choose at each information set is defined as the expectation of player 1's beliefs about player 2. This gives: $v_1 = E_2[q_1|h^3 \cup h^5]$, and $v_2 = E_2[q_2|h^4 \cup h^6]$. Player 1 can choose to cooperate or defect at node h^0. Let player 2's belief that player 1 will cooperate be $z_1 = P_2(1\, choses\, C|h^0)$. Player 1's belief about player 2's belief that player 1 will cooperate is defined as $w_1 = E_1[z_1|h^0]$.

Player 2 only observes the results of nature. Player 2's evaluation of the kindness of player 1 depends on the belief about what node she is currently at. If player 2 observes cooperation, then the probability that player 2 believes she is at node h^3 is $P(h^3|2\, observes\, C) = \frac{z_1 \cdot \varepsilon_1}{z_1 \cdot \varepsilon_1 + (1-z_1) \cdot \varepsilon_2}$ via Bayes rule. Similarly, $P(h^5|2\, observes\, C) = \frac{(1-z_1) \cdot \varepsilon_2}{z_1 \cdot \varepsilon_1 + (1-z_1) \cdot \varepsilon_2}$. If player 2 observes defection, then $P(h^4|2\, observes\, D) = \frac{z_1 \cdot (1-\varepsilon_1)}{z_1 \cdot (1-\varepsilon_1) + (1-z_1) \cdot (1-\varepsilon_2)}$ and $P(h^6|2\, observes\, D) = \frac{(1-z_1) \cdot (1-\varepsilon_2)}{z_1 \cdot (1-\varepsilon_1) + (1-z_1) \cdot (1-\varepsilon_2)}$. Since the SRE concept requires that initial beliefs be correct, it follows that player 2 knows in equilibrium what player 1 chooses.

Player 2's belief about the kindness of player 1 is $\phi_{212} = (z_1 - \frac{1}{2})(1-2\theta)(2 - v_1 + v_2)$. No matter the history, the kindness of player 2 towards player 1 will always be $k_{21} = 1$ if player 2 cooperates and $k_{21} = -1$ if player 2 defects. If at nodes h^3 and h^5, then $D_{21} = -3$ if player 2 defects and zero otherwise. If at nodes h^4 and h^6, then if player 1 cooperates $D_{21} = -3$ and is equal to zero otherwise.

Suppose that in equilibrium player 1 cooperates, then player 2 knows this. Since $z_1 = w_1 = 1$, player 2 knows that if nature cooperates then she is at node $h^3$and if nature defects then she is at node h^4. If nature cooperates, then player 2 will cooperate if $(1-2\theta)(2 - v_1 + v_2)\rho_2(1-\gamma_2) + 3\gamma_2\rho_2 > 1$. If nature defects, then player 2 will cooperate if $(1-2\theta)(2 - v_1 + v_2)\rho_2(1-\gamma_2) > 1 + 3\gamma_2\rho_2$.

For 1(a), $v_1 = 1$ and $v_2 = 0$. For this to be an equilibrium it must be that $(1-2\theta)\rho_2(1-\gamma_2) +$

$3\gamma_2\rho_2 > 1$ and $(1-2\theta)\rho_2(1-\gamma_2) < 1+3\gamma_2\rho_2$. This hold under two conditions. In the first case, if $0 < \gamma_2 < \frac{1-2\theta}{4-2\theta}$, and $\frac{1}{1-2\theta+\gamma_2(2+2\theta)} < \rho_2 < \frac{1}{1-2\theta-\gamma_2(4-2\theta)}$, then player 2 will cooperate if nature cooperates and defect if nature defects. In the second case, if$\gamma_2 > \frac{1-2\theta}{4-2\theta}$ and $\rho_2 > \frac{1}{1-2\theta+\gamma_2(2+2\theta)}$, then player 2 will cooperate if nature cooperates and defect if nature defects.

For 1(b), $v_1 = v_2 = 1$. This is possible if $(1-2\theta)(2)\rho_2(1-\gamma_2)+3\gamma_2\rho_2 > 1$ and $(1-2\theta)(2)\rho_2(1-\gamma_2) > 1+3\gamma_2\rho_2$. If $\gamma_2 < \frac{2-4\theta}{5-4\theta}$ and$\rho_2 > \frac{1}{2-4\theta-\gamma_2(5-4\theta)}$, then both conditions will be satisfied. Player 1 will cooperate no matter the results of nature.

For 1(c), $v_1 = v_2 = 0$. This is possible if both $(1-2\theta)(2)\rho_2(1-\gamma_2)+3\gamma_2\rho_2 < 1$ and $(1-2\theta)(2)\rho_2(1-\gamma_2) < 1+3\gamma_2\rho_2$. If $\gamma_2 \geq 0$ and $\rho_2 < \frac{1}{2-4\theta-\gamma_2(1-4\theta)}$, then player 2 will always defect.

If player 1 defects, then player 2's belief about the kindness of player 1 is$\phi_{212} = -\frac{1}{2}(1-2\theta)(2-v_1+v_2)$. For 2(a), $v_1 = 1$ and $v_2 = 0$. This implies that $-(1-2\theta)\rho_2(1-\gamma_2)+3\gamma_2\rho_2 > 1$ and $(1-2\theta)(2)\rho_2(1-\gamma_2)+1+3\gamma_2\rho_2 > 0$. These conditions will hold if $\gamma_2 > \frac{1-2\theta}{4-2\theta}$ and $\rho_2 > \frac{1}{\gamma_2(4-2\theta)-1+2\theta}$.

For 2(b), $v_1 = v_2 = 0$. This is possible if both $-(1-2\theta)(2)\rho_2(1-\gamma_2)+3\gamma_2\rho_2 < 1$ and $-(1-2\theta)(2)\rho_2(1-\gamma_2) < 1+3\gamma_2\rho_2$. These conditions will hold if$\gamma_2 > \frac{2-4\theta}{5-4\theta}$ and $\rho_2 < \frac{1}{\gamma_2(5-4\theta)-2+4\theta}$ or $\gamma_2 < \frac{2-4\theta}{5-4\theta}$ and$\rho_2 \geq 0$. □

To understand the equilibrium predictions when players are purely reciprocal, assume that $\gamma_i = 0$. With pure reciprocity, player 2 ignores the results of nature. As a consequence, player 2 will choose to cooperate based on the equilibrium beliefs about what player 1 chose. If player 1 cooperates with probability one, then player 1 is being kind towards player 2. Even if player 2 observes defection by nature, player 2 knows that player 1 cooperated and player 1 is still viewed as kind.

The control that player 1 has still matters. When player 1 has more control, the value of ρ_2 needed for player 2 to cooperate can be smaller all other things equal. This suggests that cooperation should be higher when player 1 has more control. Notice however that cooperation in pure strategies is still possible even when control is low. This differs from the perfect information game.

If players are purely inequity averse, then $\gamma_i = 1$. The equilibrium predictions for a player 2 with pure inequity aversion are the same for the perfect or imperfect information games. This makes sense because inequity aversion is only outcome based, and player 's1 intended choice does not influence player 2's fairness judgments.

With mixed concerns, the potential equilibrium in 1(a) gives that player 2 will cooperate if nature cooperates and defect if nature defects. This equilibrium can occur if player 2 is strongly concerned about inequity aversion. Notice, however, that the equilibrium is also possible for a player 2 that cares a great deal about reciprocity. For certain ranges of ρ_2, a player that is highly reciprocal will behave as if they are concerned about inequity aversion. This suggests that as control changes the types that players appear to be could change as well. As a result, it is possible that some players could behave inequity averse, self-interested, or reciprocal depending upon player 1's level of control. In 1(b), player 2 will cooperate regardless of nature's choice. In equilibrium. player 2 knows that player 1 cooperated and cooperation by player 1 is viewed as kind. This kindness is enough for players that are highly concerned about reciprocity to cooperate even if nature defects.

There are a large number of potential equilibria that can occur for player 1 due to self-fulfilling expectations. The focus of this paper on second mover behavior. In interest of space, equilibrium predictions for player 1 are left to the Supplemental Appendix.

Perfect versus Imperfect Information

Both the perfect and imperfect information games can be used to test the predictions of the fairness models explored in this paper. Predictions from pure self-interest and pure inequity aversion are the same no matter the information. As a result of these models, changes in the information about what player 1 chose should not be relevant for equilibrium behavior.

With pure reciprocity, equilibrium behavior could differ depending upon the information available to player 2. In the perfect information game, pure strategy cooperation by player 2 only occurs if control is high. However, in the imperfect information game, cooperation is still possible when control is low. Even when control is high, the concern for reciprocity needs to be much higher when information is perfect compared to the imperfect information game in order for cooperation to be possible. As a result, if subjects are motivated by reciprocity, then cooperation should be higher when information is imperfect compared to when the information is perfect.

In the mixed concerns model, when control is high in the perfect information game, it is possible to have an equilibrium in which player 2 cooperates if player 1 cooperates and defects if player 1 defects. However, when control is low this equilibrium no longer exists. This is not the case with the imperfect information game. When control is low it is still possible for player 2 to cooperate if player 1 cooperates and defect if player 1 defects. Even when control is high, the range of values for both ρ_2 and γ_2 that lead player 2 to cooperate is largest in the imperfect information game. Thus, given that player 1 cooperates, cooperation by player 2 in the imperfect information game should be higher than in the perfect information game.

A.5 Additional Data Analysis and Robustness Checks

Table A.1 presents the the raw empathy and perspective taking scores by treatment and condition. A total of six subjects failed to complete the empathic concern questions and a total of four subjects failed to complete the perspective taking questions. A total of eight subjects failed to complete both the empathic concern and perspective taking questions. In the Known treatment, there was no statistical difference in empathic concern scores in subjects that received the high control first compared to low control (Wilcoxon rank-sum test, N=121, z=-1.10, p=0.27). Results are similar in the Uncertain treatment (Wilcoxon rank-sum test, N=118, z=1.59, p=0.11). When subjects received the High Control condition first, there was no significant difference in empathic concern in the Known treatment compared to the Uncertain treatment (Wilcoxon rank-sum test, N=122, z=-1.58, p=0.11). Similary results occur when subjects received the Low Control condition first (Wilcoxon rank sum test, N=117, z=1.14, p=0.25).

Table A.1: Empathy and Perspective Taking Summary Statistics

	High Control First		Low Control First		
	Known	Uncertain	Known	Uncertain	Total
Average:					
Empathic Concern	24.7	25.7	25.1	24.6	25.1
Number of Subjects	60	62	61	56	239
Perspective Taking	26.3	26.0	25.5	25.6	25.9
Number of Subjects	61	63	61	57	242

In the Known treatment, perspective taking ability did not significantly differ between subjects who received the High Control condition first compared to those who received the Low Control condition first (Wilcoxon rank-sum test, N=122, z=1.26, p=0.21). Similarly results hold for the Uncertain treatment (Wilcoxon rank-sum test, N=120, z=0.71, p=0.48). When subjects received the High Control condition first, there was no significant difference between the Known and Uncertain treatments in perspective taking (Wilcoxon rank-sum test, N=124, z=0.419, p=0.68). Similar results are found when subjects received the Low Control condition first (Wilcoxon rank-sum test, N=118, z=-0.06, p=0.95).

Table A.2 presents robustness checks for the results from table 1.3 in the main paper. The analysis is repeated except for the addition of one period lagged variables for the the possible paths of play. The lagged variables are only significant at the 10% in the Known treatment and High Control condition. The hypothesis tests for inequity aversion and reciprocity are similar to the table in the main paper. Table A.3 reports fixed effects logit regressions as a further robustness check for table 1.3 from the main paper. The table gives odds ratios on the probability that the second mover will cooperate given the path of play. The hypothesis tests are similar to table 1.3 as well.

Table A.2: Second Mover Cooperation by Treatment with lagged variables

	Known		Uncertain	
	High Control	Low Control	High Control	Low Control
First Mover and Computer cooperated	1.82***	2.18***	1.45***	1.62***
	(0.28)	(0.34)	(0.23)	(0.31)
First Mover cooperated and Computer defected		−0.14	0.04	0.42
		(0.49)	(0.65)	(0.37)
First Mover defected and Computer Cooperated	1.11**	1.75***	1.76***	1.69***
	(0.34)	(0.34)	(0.34)	(0.31)
First Mover and Computer cooperated lag	0.49+	−0.19	−0.06	0.11
	(0.28)	(0.26)	(0.24)	(0.28)
First Mover cooperated and Computer defected lag	−0.29	0.40	−0.10	0.48
	(0.52)	(0.34)	(0.53)	(0.32)
First Mover defected and Computer cooperated lag	0.57+	0.19	0.12	0.01
	(0.33)	(0.29)	(0.34)	(0.27)
Low Control First	0.15	0.66	−0.67	1.67*
	(0.64)	(0.72)	(0.60)	(0.79)
Female	−0.22	0.12	−0.63+	−0.26
	(0.38)	(0.43)	(0.37)	(0.55)
Intercept	−1.86**	−3.28***	−0.70	−2.82***
	(0.60)	(0.67)	(0.52)	(0.75)
N	556	604	562	563
ρ	0.56	0.64	0.51	0.72
Model χ^2	57.85	57.64	59.98	58.48
Hypothesis Tests				
Inequity Aversion (Prob> $\chi^2(1)$)	0.04*	0.08+	0.35	0.79
Reciprocity (Prob> $\chi^2(1)$)	0.00***	0.00***	0.03*	0.00***

Cluster robust standard errors by subject in parentheses. Hypothesis for Inequity Aversion is that cooperation given computer cooperated is the same regardless of first mover's choice. Hypothesis for Reciprocity is that cooperation given first mover cooperated is the same regardless of computer's choice.
Results are from random-effects probit regressions with round fixed effects.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

One worry is that there is multicollinearity between the empathic concern and perspective taking variables. Table A.4 repeats the analysis from table 1.4 in the main paper restricting the regressions to only include empathic concern or perspective taking. Comparing to table 1.4 all the regressions have similar results except in a few cases. In column (2) of table A.4 the coefficient for empathic

Table A.3: Second Mover Cooperation by Treatment- Logit Regressions

	Known		Uncertain	
	High Control	Low Control	High Control	Low Control
First Mover and Computer cooperated	3.06***	3.25***	2.84***	2.64***
	(0.49)	(0.53)	(0.47)	(0.53)
First Mover cooperated and Computer defected	−0.70	−0.68	0.13	0.67
	(1.29)	(0.88)	(1.15)	(0.61)
First Mover defected and Computer cooperated	1.27*	2.47***	3.03***	2.74***
	(0.56)	(0.53)	(0.62)	(0.50)
N	300	320	360	310
Model χ^2	97.70	107.69	101.39	95.06
Hypothesis Tests				
Inequity Aversion (Prob> $\chi^2(1)$)	0.01**	0.08+	0.76	0.82
Reciprocity (Prob> $\chi^2(1)$)	0.01**	0.00***	0.02*	0.00***

Cluster robust standard errors by subject in parentheses. Hypothesis for Inequity Aversion is that cooperation given computer cooperated is the same regardless of first mover's choice. Hypothesis for Reciprocity is that cooperation given first mover cooperated is the same regardless of computer's choice.
Results are from fixed effects logit regressions with round fixed effects.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

concern is no longer significant at the 10% level. However, table A.5 column(1) shows that empathic concern is still significant when the regression is restricted to cases when the first mover cooperated. In column (4) of table A.4, unlike the regression results in the main paper perspective taking is signifigant at the 5% level. The results for perspective taking are similar in table A.5 to table 1.5 in the main paper. Table A.6 and A.7 examine empathic concern and perspective taking looking at all 20 rounds finding similar results.

Table A.4: Second Mover Cooperation with Empathy and Perspective Taking Robustness(First 10 Rounds)

	All Treatments	Known	All Treatments		Known	
	(1)	(2)	(3)	(4)	(5)	(6)
	Cooperation	Cooperation	Cooperation	Cooperation	Cooperation	Cooperation
Empathic Concern	0.32^{**}	0.31				
	(0.12)	(0.21)				
Perspective Taking			0.18	0.31^{*}	−0.03	0.32
			(0.12)	(0.13)	(0.21)	(0.25)
Low Control	0.55^{**}	0.53	0.51^{*}	0.52^{*}	0.48	0.41
	(0.21)	(0.34)	(0.22)	(0.22)	(0.34)	(0.34)
Low Control X Perspective Taking				−0.32		-0.89^{**}
				(0.26)		(0.34)
Uncertain	0.51^{*}		0.45^{*}	0.45^{*}		
	(0.21)		(0.21)	(0.21)		
First Mover and Computer cooperated	1.33^{***}	1.43^{***}	1.29^{***}	1.29^{***}	1.36^{***}	1.35^{***}
	(0.18)	(0.28)	(0.17)	(0.17)	(0.27)	(0.27)
First Mover cooperated and Computer defected	0.04	−0.36	0.10	0.10	−0.43	−0.44
	(0.24)	(0.49)	(0.24)	(0.24)	(0.48)	(0.50)
First Mover defected and Computer Cooperated	1.23^{***}	1.15^{***}	1.19^{***}	1.19^{***}	1.09^{***}	1.11^{***}
	(0.19)	(0.28)	(0.18)	(0.18)	(0.26)	(0.26)
Female	-0.36^{+}	−0.15	−0.19	−0.17	0.05	0.17
	(0.21)	(0.33)	(0.21)	(0.21)	(0.34)	(0.32)
Intercept	-1.52^{***}	-1.62^{***}	-1.54^{***}	-1.58^{***}	-1.65^{***}	-1.73^{***}
	(0.28)	(0.45)	(0.29)	(0.29)	(0.47)	(0.46)
N	1210	610	1220	1220	620	620
ρ	0.43^{***}	0.51^{***}	0.44^{***}	0.44^{***}	0.51^{***}	0.48^{***}
Model χ^2	138.02	63.92	135.91	136.92	64.22	65.79

Results are from random effects probit regressions with round fixed effects. Regressions are from first 10 rounds. Cluster robust standard at the subject level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Second Mover Conditional Cooperation with Empathy and Perspective Taking Robustness

	(1) First Mover Cooperated Cooperation	(2) First Mover Defected Cooperation	(3) First Mover Cooperated Cooperation	(4) First Mover Defected Cooperation
Empathic Concern	0.51^{+}	0.15		
	(0.29)	(0.20)		
Perspective Taking			0.36	0.23
			(0.40)	(0.19)
Low Control	0.61	0.47	0.56	0.30
	(0.54)	(0.31)	(0.54)	(0.32)
Low Control X Perspective Taking			-1.01^{+}	-0.72^{*}
			(0.52)	(0.33)
Computer cooperated	2.15^{*}	1.08^{***}	2.16^{*}	1.10^{***}
	(0.88)	(0.26)	(0.89)	(0.24)
Female	-0.25	-0.12	0.17	0.10
	(0.52)	(0.36)	(0.49)	(0.33)
Intercept	-2.76^{*}	-0.90^{+}	-3.05^{**}	-1.06^{*}
	(1.11)	(0.49)	(1.17)	(0.48)
N	209	306	210	313
ρ	0.67^{***}	0.37^{***}	0.65^{***}	0.37^{***}
Model χ^2	20.87	27.59	20.44	28.93

Results are from random effects probit regressions with round fixed effects. Cluster robust standard errors at the subject level in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Second Mover Conditional Cooperation in Known Treatment with Empathy- Robustness

	High Control		Low Control	
	First Mover Cooperated	First Mover Defected	First Mover Cooperated	First Mover Defected
Empathic Concern	0.81*	0.18	0.86*	−0.06
	(0.39)	(0.24)	(0.36)	(0.33)
Computer Cooperated	1.76	1.50**	2.58***	2.10***
	(1.15)	(0.52)	(0.70)	(0.59)
Low Control First	0.25	−0.75	1.02	2.33*
	(0.89)	(0.70)	(1.09)	(1.17)
Female	−0.49	−0.32	−0.21	0.17
	(0.65)	(0.47)	(0.54)	(0.60)
Intercept	−2.36+	−0.67	−3.66*	−3.76***
	(1.38)	(0.63)	(1.44)	(1.03)
N	135	231	236	266
ρ	0.72	0.52	0.69	0.71
Model χ^2	11.68	19.40	25.65	21.80

Regressions are from random effects probit regressions with round fixed effects. Cluster robust standard errors at the subject level in parentheses. Regressions include all 20 rounds.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: Second Mover Conditional Cooperation in Known Treatment with Perspective Taking-Robustness

	High Control		Low Control	
	First Mover Cooperated	First Mover Defected	First Mover Cooperated	First Mover Defected
Perspective Taking	0.39	0.23	0.56	0.34
	(0.48)	(0.22)	(0.43)	(0.40)
Low Control First	0.29	-0.53	0.66	2.09^{+}
	(0.91)	(0.60)	(0.90)	(1.07)
Low Control First X Perspective Taking	-0.87	-0.60^{+}	-1.36^{*}	-1.17^{+}
	(0.84)	(0.35)	(0.55)	(0.61)
Computer Cooperated	1.79	1.20^{***}	2.50^{***}	1.96^{***}
	(1.28)	(0.35)	(0.69)	(0.50)
Female	0.02	-0.03	0.44	0.38
	(0.64)	(0.36)	(0.48)	(0.53)
Intercept	-2.86^{+}	-0.93^{+}	-3.63^{**}	-3.69^{***}
	(1.61)	(0.56)	(1.27)	(0.94)
N	136	282	239	288
ρ	0.76	0.43	0.65	0.68
Model χ^2	9.76	23.59	25.37	22.57

Regressions are from random effects probit regressions with round fixed effects. Cluster robust standard errors at the subject level in parentheses. Regressions include all 20 rounds.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.6 Experiment Instructions

The following screenshots are the instructions viewed by subjects in the experiment for the treatments that received the High Control condition in the first ten rounds. When subjects received the Low Control condition first, subjects viewed the exact same instructions except that the 10% reversal probability in the first part of the instructions was changed to 40%. For the second part, the instructions remained the same except that the 40% reversal probability was changed to 10%.

Part I Instructions

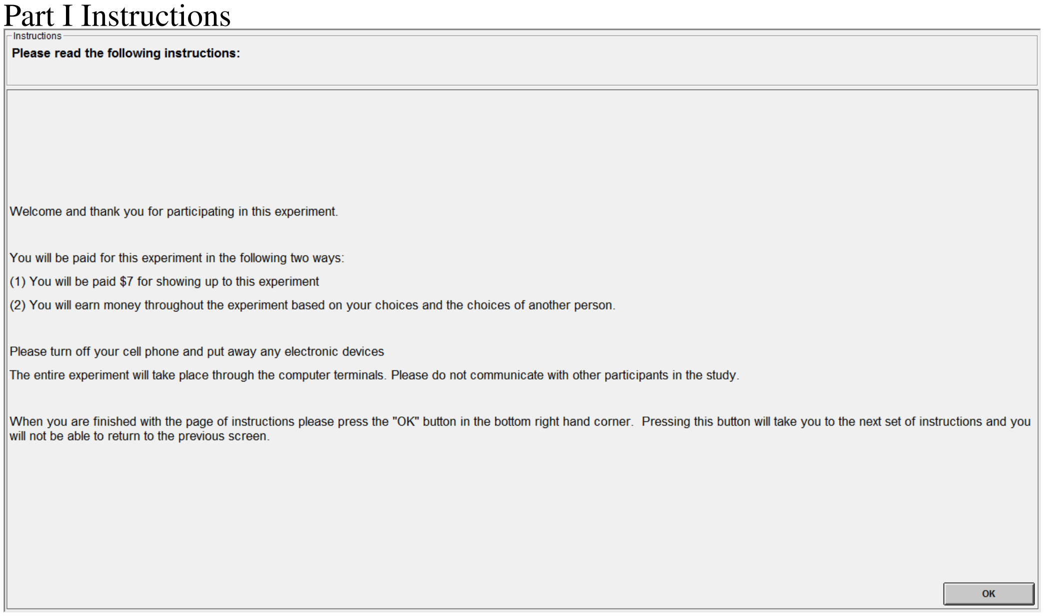


Figure A.1: Experimental Instructions

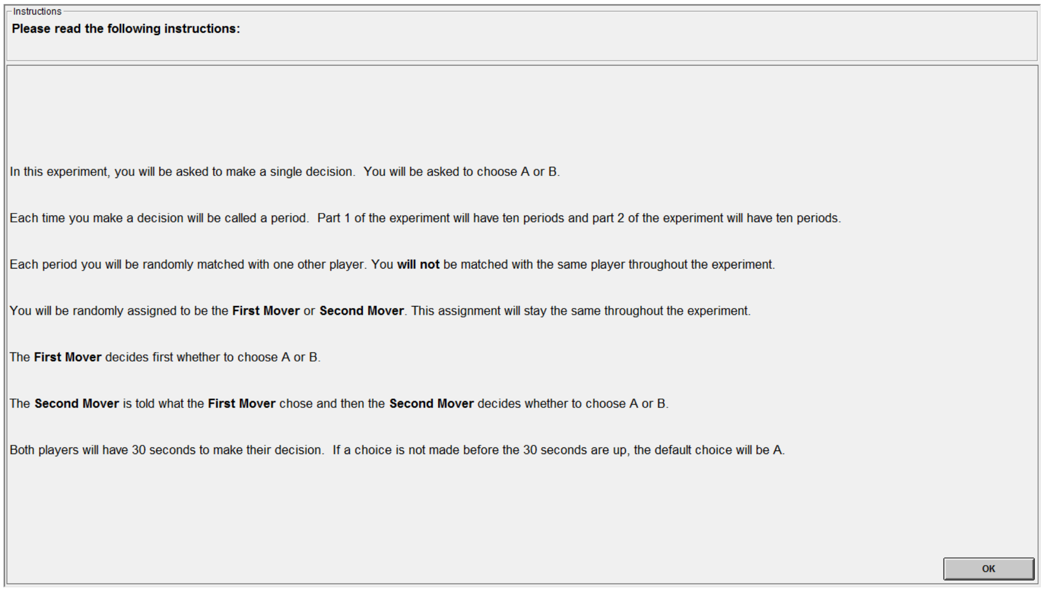


Figure A.2: Experimental Instructions- continued

Figure A.3: Experimental Instructions- continued

Instructions

Please read the following instructions:

The **Second Mover** will be told what the **First Mover** chose and what the computer chose for the **First Mover**.

For example, if the **First Mover** selected B and the computer changed the choice to A, then the **Second Mover** will be told that the **First Mover** chose B and the computer changed the choice to A.

If the **First Mover** selected A and the computer did not change the choice, then the **Second Mover** will be told that the **First Mover** chose A and the computer chose A.

The computer will not change the **Second Mover's** choice.
If the **Second Mover** chooses A, then the choice will be A.
If the **Second Mover** chooses B, then the choice will be B.

Only the **First Mover's** choice has a chance to be changed.

OK

Figure A.4: Experimental Instructions- Known Treatment Only

Instructions

Please read the following instructions:

The **Second Mover** will only be told what the computer chose for the **First Mover**. The **Second Mover** will not be told what the **First Mover** chose.

For example, if the **First Mover** selected B and the computer changed the choice to A, then the **Second Mover** will be told that the computer chose A for the **First Mover**.

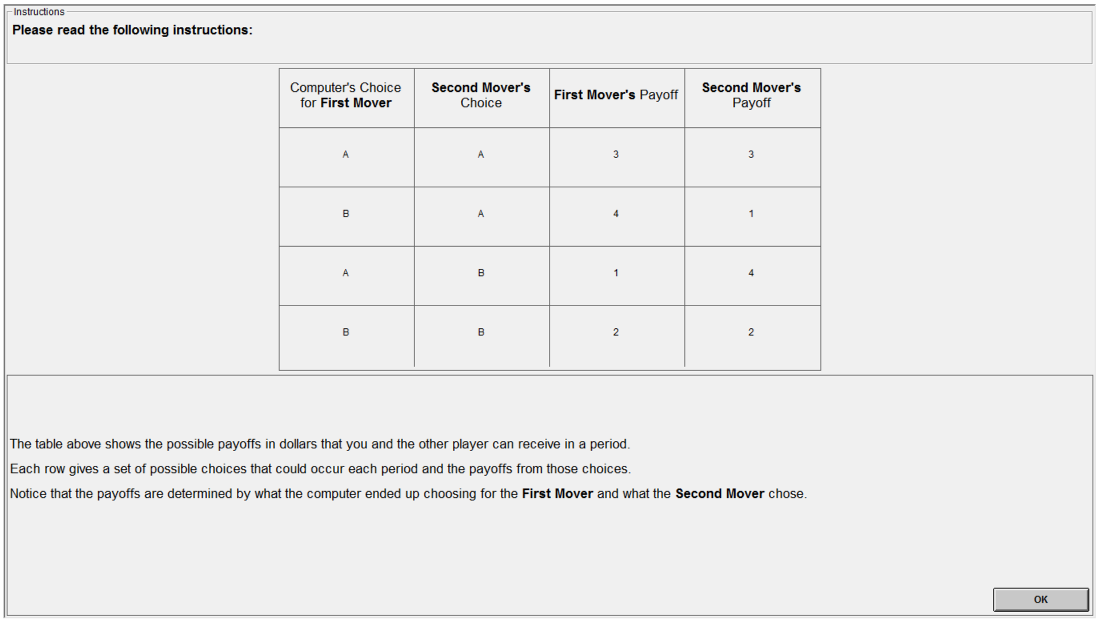
If the **First Mover** selected A and the computer did not change the choice, then the **Second Mover** will be told that the computer chose A.

The computer will not change the **Second Mover's** choice.
If the **Second Mover** chooses A, then the choice will be A.
If the **Second Mover** chooses B, then the choice will be B.

Only the **First Mover's** choice has a chance to be changed.

OK

Figure A.5: Experimental Instructions- Uncertain Treatment Only



Instructions

Please read the following instructions:

Computer's Choice for **First Mover**	**Second Mover's** Choice	**First Mover's** Payoff	**Second Mover's** Payoff
A	A	3	3
B	A	4	1
A	B	1	4
B	B	2	2

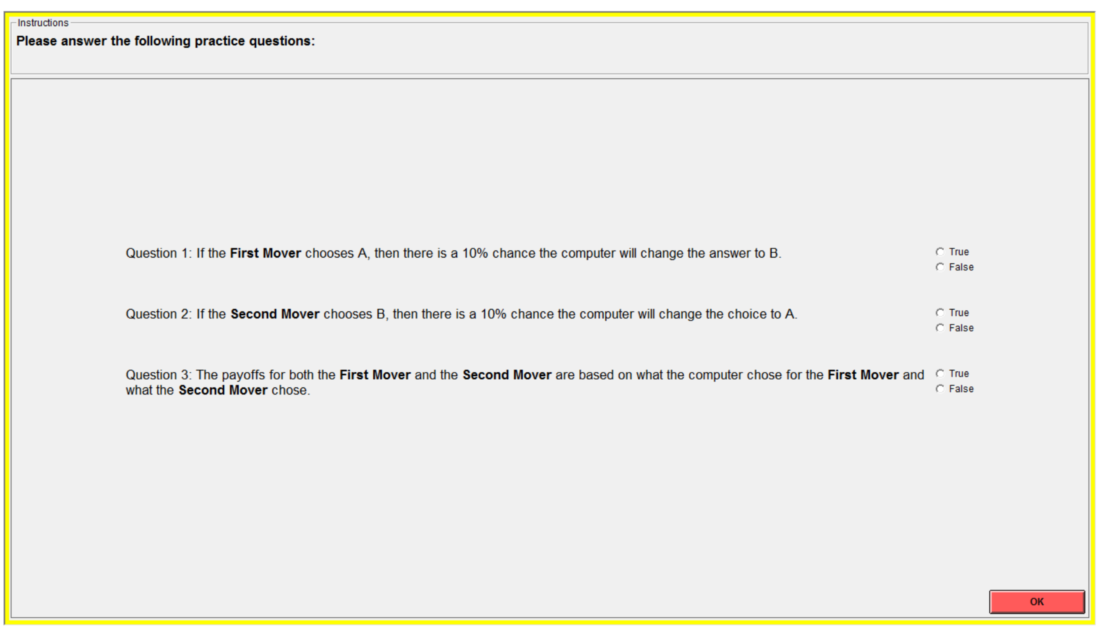
The table above shows the possible payoffs in dollars that you and the other player can receive in a period.

Each row gives a set of possible choices that could occur each period and the payoffs from those choices.

Notice that the payoffs are determined by what the computer ended up choosing for the **First Mover** and what the **Second Mover** chose.

OK

Figure A.6: Experimental Instructions- Payoff Table



Instructions

Please answer the following practice questions:

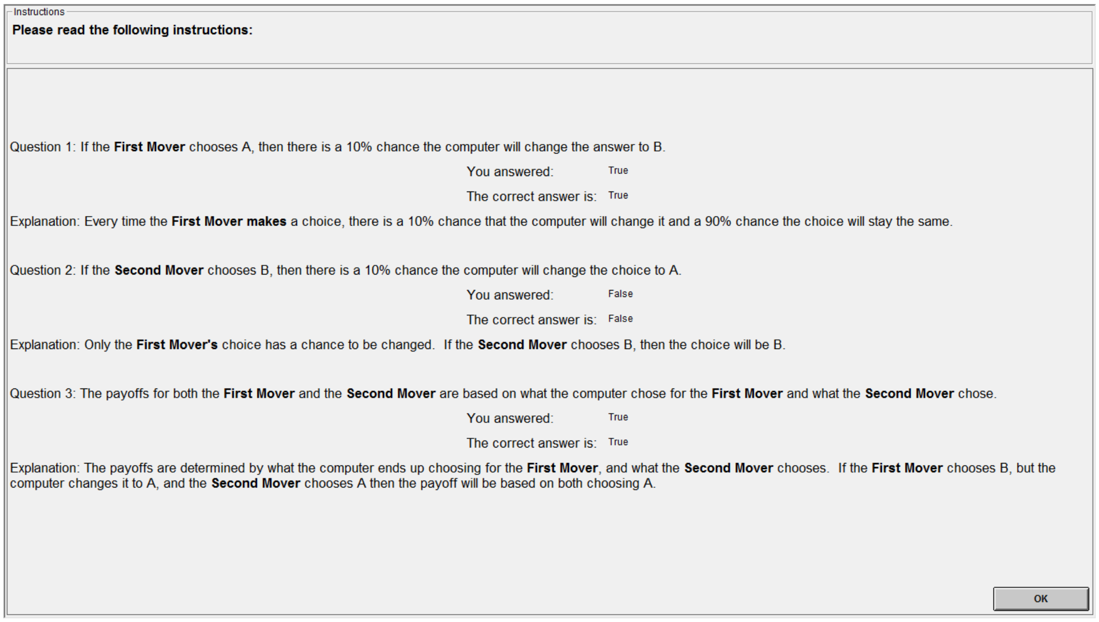
Question 1: If the **First Mover** chooses A, then there is a 10% chance the computer will change the answer to B. True / False

Question 2: If the **Second Mover** chooses B, then there is a 10% chance the computer will change the choice to A. True / False

Question 3: The payoffs for both the **First Mover** and the **Second Mover** are based on what the computer chose for the **First Mover** and what the **Second Mover** chose. True / False

OK

Figure A.7: Experimental Instructions- Comprehension Questions



Instructions

Please read the following instructions:

Question 1: If the **First Mover** chooses A, then there is a 10% chance the computer will change the answer to B.

You answered: True

The correct answer is: True

Explanation: Every time the **First Mover makes** a choice, there is a 10% chance that the computer will change it and a 90% chance the choice will stay the same.

Question 2: If the **Second Mover** chooses B, then there is a 10% chance the computer will change the choice to A.

You answered: False

The correct answer is: False

Explanation: Only the **First Mover's** choice has a chance to be changed. If the **Second Mover** chooses B, then the choice will be B.

Question 3: The payoffs for both the **First Mover** and the **Second Mover** are based on what the computer chose for the **First Mover** and what the **Second Mover** chose.

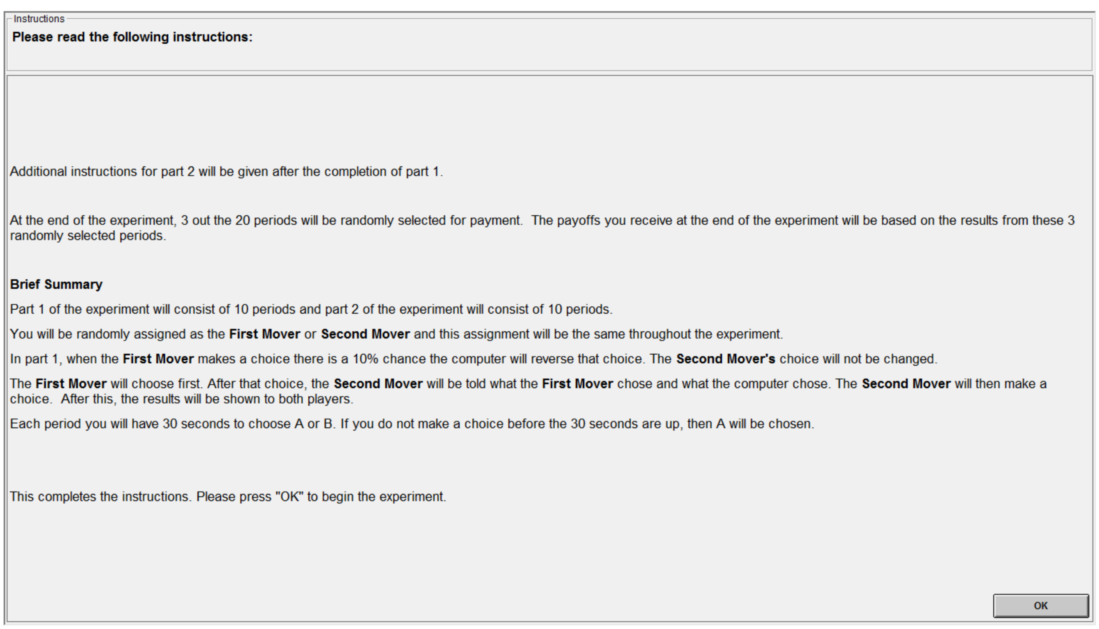
You answered: True

The correct answer is: True

Explanation: The payoffs are determined by what the computer ends up choosing for the **First Mover**, and what the **Second Mover** chooses. If the **First Mover** chooses B, but the computer changes it to A, and the **Second Mover** chooses A then the payoff will be based on both choosing A.

OK

Figure A.8: Experimental Instructions- Comprehension Answers



Instructions

Please read the following instructions:

Additional instructions for part 2 will be given after the completion of part 1.

At the end of the experiment, 3 out the 20 periods will be randomly selected for payment. The payoffs you receive at the end of the experiment will be based on the results from these 3 randomly selected periods.

Brief Summary

Part 1 of the experiment will consist of 10 periods and part 2 of the experiment will consist of 10 periods.

You will be randomly assigned as the **First Mover** or **Second Mover** and this assignment will be the same throughout the experiment.

In part 1, when the **First Mover** makes a choice there is a 10% chance the computer will reverse that choice. The **Second Mover's** choice will not be changed.

The **First Mover** will choose first. After that choice, the **Second Mover** will be told what the **First Mover** chose and what the computer chose. The **Second Mover** will then make a choice. After this, the results will be shown to both players.

Each period you will have 30 seconds to choose A or B. If you do not make a choice before the 30 seconds are up, then A will be chosen.

This completes the instructions. Please press "OK" to begin the experiment.

OK

Figure A.9: Experimental Instructions- Known Treatment Only

Instructions

Please read the following instructions:

Additional instructions for part 2 will be given after the completion of part 1.

At the end of the experiment, 3 out the 20 periods will be randomly selected for payment. The payoffs you receive at the end of the experiment will be based on the results from these 3 randomly selected periods.

Brief Summary

Part 1 of the experiment will consist of 10 periods and part 2 of the experiment will consist of 10 periods.

You will be randomly assigned as the **First Mover** or **Second Mover** and this assignment will be the same throughout the experiment.

In part 1, when the **First Mover** makes a choice there is a 10% chance the computer will reverse that choice. The **Second Mover's** choice will not be changed.

The **First Mover** will choose first. After that choice, the **Second Mover** will only be told what the computer chose. The **Second Mover** will then make a choice. After this, the results will be shown to both players.

Each period you will have 30 seconds to choose A or B. If you do not make a choice before the 30 seconds are up, then A will be chosen.

This completes the instructions. Please press "OK" to begin the experiment.

OK

Figure A.10: Experimental Instructions- Uncertain Treatment Only

Instructions

Please read the following instructions:

In part 2 of the experiment, there is now a 40% chance that the choice the First Mover makes will be reversed by the computer.

For example, if the **First Mover** chooses A, then there is a 60% chance that the choice stays as A and a 40% chance that the choice changes to B.

If the **First Mover** chooses B, then there is a 60% chance the choice stays as B and a 40% chance the choice changes to A.

The **Second Mover** will still be told what the **First Mover** chose and what the computer chose.

OK

Figure A.11: Experimental Instructions Part II- Known Treatment Only

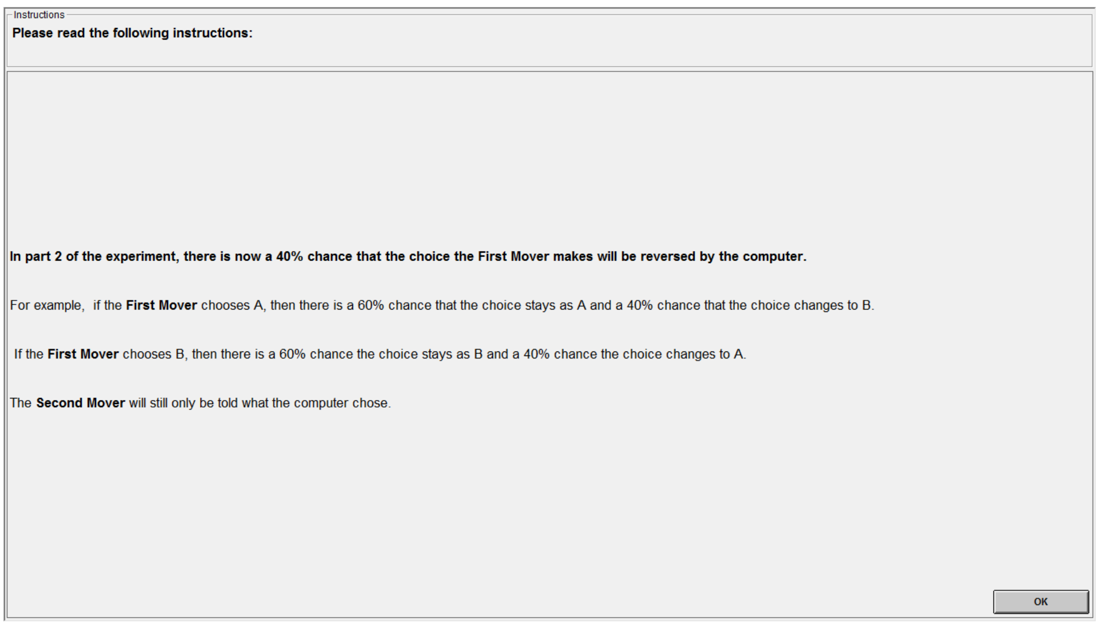


Figure A.12: Experimental Instructions Part II- Uncertain Treatment Only

B Appendix

B.1 Norm-based utility predictions

Here I present the equilibrium predictions from the norm based utility function introduced in Section 3 of the paper. To simplify the analysis, the alternative offers of the mini-ultimatum games for the proposer is denoted as x_1^A and for the responder is denoted as x_2^A. The norm $\theta_i(h)$ can differ based on the information set h. Let h_0 denote the information set when the proposer makes their offer, h_1 be the information set that is reached if the proposer offers (8,2), and h_2 be the information set that is reached if the proposer offers the alternative (x_1^A, x_2^A).

Proposition 8. *The potential subgame perfect equilibrium for the responder are:*

1. If the proposer offers (8,2):

   (a) With strong punishment:

      i. If $\theta_2 > \frac{1}{5}$ and $\gamma_2 > \frac{2}{g(10\theta_2(h_1)-2)}$, then the responder will reject the (8,2) offer.

      ii. If $\theta_2 < \frac{1}{5}$ or $\gamma_2 < \frac{2}{g(10\theta_2(h_1)-2)}$, then the responder will accept the (8,2) offer.

   (b) With weak punishment:

      i. If $\theta_2 > \frac{1}{4}$ and $\gamma_2 > \frac{1}{g(10\theta_2(h_1)-2)-g(2\theta_2)}$, then the responder will reject the (8,2) offer.

      ii. If $\theta_2 < \frac{1}{4}$ or $\gamma_2 < \frac{1}{g(10\theta_2(h_1)-2)-g(2\theta_2(h_1))}$, then the responder will accept the (8,2) offer.

2. If the proposer offers (x_1^A, x_2^A):

   (a) With strong punishment:

      i. If $\theta_2 > \frac{x_2^A}{10}$ and $\gamma_2 > \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)}$, then the responder will reject the (x_1^A, x_2^A) offer.

      ii. If $\theta_2 < \frac{\pi_2^A}{10}$ or $\gamma_2 < \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)}$,then the responder will accept the (x_1^A, x_2^A) offer.

(b) With weak punishment:

i. If $\theta_2 > \frac{\pi_2^A}{8}$ and $\gamma_2 > \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)-g(2\theta_2)}$, then the responder will reject the (x_1^A,x_2^A) offer.

ii. If $\theta_2 < \frac{\pi_2^A}{8}$ or $\gamma_2 < \frac{\pi_2^A}{g(10\theta_2(h_2)-\pi_2^A)-g(2\theta_2(h_2))}$, then the responder will accept the (x_1^A,x_2^A) offer.

Proof. With strong punishment, the utility of rejection is zero. This occurs because rejection reduces the entire surplus X to zero. For 1(a), the responder will reject the (8,2) offer if $2-\gamma_2 \cdot g(10\theta_2(h_1)-2) < 0$. This holds if $\theta_2 > \frac{1}{5}$ and $\gamma_2 > \frac{2}{g(10\theta_2(h_1)-2)}$. The responder will accept the (8,2) offer if $2-\gamma_2 \cdot g(10\theta_2(h_1)-2) > 0$. This will hold if $\theta_2 < \frac{1}{5}$ or $\gamma_2 < \frac{2}{g(10\theta_2(h_1)-2)}$. With weak punishment, the utility of rejection is no longer zero. In this case, the surplus is reduced to 2. For 1(b), the responder will reject the (8,2) offer if $2-\gamma_2 \cdot g(10\theta_2(h_1)-2) < -\gamma_2 \cdot g(2\theta_2(h_1))$. This will hold if $\theta_2 > \frac{1}{4}$ and $\gamma_2 > \frac{1}{g(10\theta_2(h_1)-2)-g(2\theta_2)}$. The responder will accept the (8,2) offer if $2-\gamma_2 \cdot g(10\theta_2(h_1)-2) < -\gamma_2 \cdot g(2\theta_2(h_1))$. This will hold if $\theta_2 < \frac{1}{4}$ or $\gamma_2 < \frac{1}{g(10\theta_2(h_1)-2)-g(2\theta_2(h_1))}$.

For 2(a), the responder will reject the (x_1^A,x_2^A) offer if $x_2^A - \gamma_2 \cdot g(10\theta_2(h_2)-x_2^A) < 0$. This holds if $\theta_2 > \frac{x_2^A}{10}$ and $\gamma_2 > \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)}$. Similarly, the responder will accept if $\theta_2 < \frac{\pi_2^A}{10}$ or $\gamma_2 < \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)}$. For 2(b), the responder will reject if $x_2^A - \gamma_2 \cdot g(10\theta_2(h_2)-x_2^A) < -\gamma_2 \cdot g(10\theta_2(h_2)$. This holds if $\theta_2 > \frac{\pi_2^A}{8}$ and $\gamma_2 > \frac{x_2^A}{g(10\theta_2(h_2)-x_2^A)-g(2\theta_2)}$. Similarly, the responder will accept if $\theta_2 < \frac{\pi_2^A}{8}$ or $\gamma_2 < \frac{\pi_2^A}{g(10\theta_2(h_2)-\pi_2^A)-g(2\theta_2(h_2))}$, □

Proposition 9. *The potential subgame perfect equilibrium for the proposer are:*

1. If the responder accepts (8,2) and the alternative (x_1^A,x_2^A)

   (a) If $x_1^A < 8$, then the proposer will offer (8,2).

   (b) If $x_1^A > 8$, then the proposer will offer (x_1^A,x_2^A).

2. With strong punishment and if the responder rejects (8,2) and accepts the alternative (x_1^A,x_2^A):

(a) If $\theta_1 > \frac{x_1^A}{10}$ and $\gamma_1 > \frac{x_1^A}{g(10\theta_1(h_0)-x_1^A)}$, then the proposer will offer (8,2).

(b) If $\theta_1 < \frac{\pi_1^A}{10}$ or $\gamma_1 < \frac{\pi_1^A}{g(10\theta_1(h_0)-\pi_1^A)}$, then the proposer will offer (x_1^A,x_2^A).

3. With weak punishment and if the responder rejects (8,2) and accepts the alternative (x_1^A,x_2^A):

   (a) If $\theta_1 > \frac{\pi_1^A}{10}$ and $\gamma_1 > \frac{x_1^A-2}{g(10\theta_1(h_0)-x_1^A)}$, then the proposer will offer (8,2).

   (b) If $\theta_1 < \frac{\pi_1^A}{10}$ or $\gamma_1 < \frac{x_1^A-2}{g(10\theta_1(h_0)-x_1^A)}$, then the proposer will offer (x_1^A,x_2^A).

4. If the responder rejects both (8,2) and (x_1^A,x_2^A), then the proposer is indifferent between the two offers and will offer (8,2) with some probability p.

Proof. For 1(a), the proposer will choose the (8,2) offer if $8-\gamma_1\cdot g(10\cdot\theta_1(h_0)-8) > x_1^A-\gamma_1\cdot g(10\cdot\theta_1(h_0)-x_1^A)$. Note that if $x_1^A < 8$, then $\cdot g(10\cdot\theta_1(h_0)-8)-g(10\cdot\theta_1(h_0)-x_1^A) < 0$ since $g()$ is a monotonically increasing function. So, if If $x_1^A \leq 8$, then the proposer will offer (8,2). Similarly reasoning shows that if $x_1^A > 8$, then the proposer will offer (x_1^A,x_2^A).

For 2(a), the proposer will select the (8,2) offer if $x_1^A-\gamma_1\cdot g(10\cdot\theta_1(h_0)-x_1^A) < 0$. This inequality will hold if $\theta_1 > \frac{x_1^A}{10}$ and $\gamma_1 > \frac{x_1^A}{g(10\theta_1(h_0)-x_1^A)}$. For 2(b), the proposer will offer (x_1^A,x_2^A) if $x_1^A-\gamma_1\cdot g(10\cdot\theta_1(h_0)-x_1^A) > 0$ which holds if $\theta_1 < \frac{\pi_1^A}{10}$ or $\gamma_1 < \frac{\pi_1^A}{g(10\theta_1(h_0)-\pi_1^A)}$.

For 3(a), the proposer will select the (8,2) offer if $x_1^A-\gamma_1\cdot g(10\cdot\theta_1(h_0)-x_1^A) < 2$. this holds if $\theta_1 > \frac{\pi_1^A}{10}$ and $\gamma_1 > \frac{x_1^A-2}{g(10\theta_1(h_0)-x_1^A)}$. Similar reasoning shows that the proposer will offer (x_1^A,x_2^A) if $\theta_1 < \frac{\pi_1^A}{10}$ or $\gamma_1 < \frac{x_1^A-2}{g(10\theta_1(h_0)-x_1^A)}$.

For 4, the proposer is indifferent between the two options as the utility the get from the two choices is the same. As a result, the proposer will offer (8,2) with some probability p. □

C Appendix

C.1 Additional Tables and Graphs

Table C.8: Summary Statistics for Emotion Variables with Statistical Tests

	(1) Proposer Surplus	(2) Responder Surplus	Wilcoxon Rank-sum Test z-score	p-value
Proposer Guilt	20.92	20.56	0.56	0.58
	(4.91)	(4.63)		
	N=50	N=48		
Proposer Shame	22.54	21.87	0.54	0.59
	(4.00)	(4.71)		
	N=50	N=47		
Proposer RMET	27.02	26.98	0.05	0.96
	(3.99)	(4.18)		
	N=50	N=49		
Responder Guilt	19.94	21.39	−1.55	0.12
	(4.99)	(4.42)		
	N=49	N=49		
Responder Shame	22.18	22.20	0.07	0.43
	(4.46)	(4.12)		
	N=49	N=49		
Responder RMET	26.58	27.78	−1.61	0.11
	(4.09)	(4.00)		
	N=50	N=49		

Wilcoxon Rank-Sum tests the hypothesis that the variable in columns (1) and (2) come from same distribution.

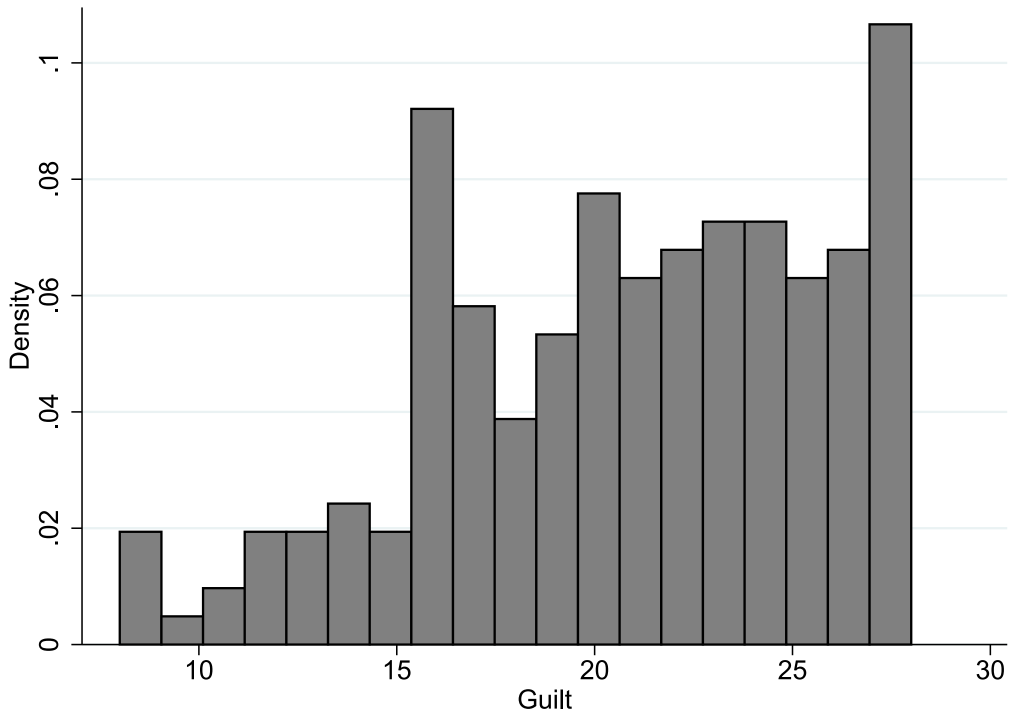


Figure C.13: Histogram of Guilt Score

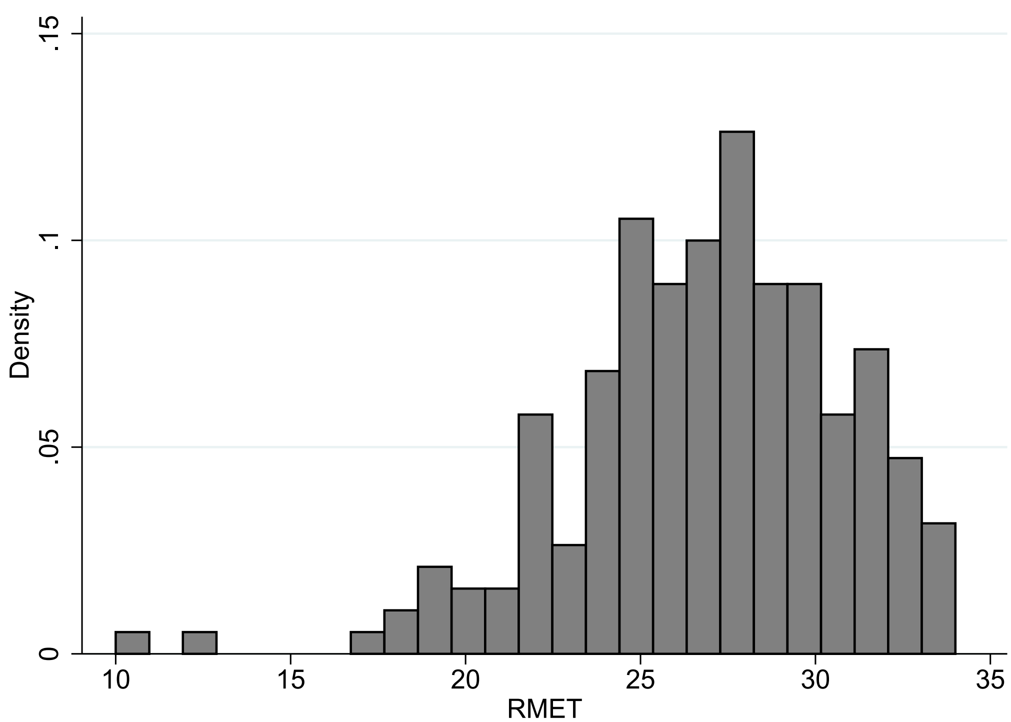


Figure C.14: Histogram of RMET Score

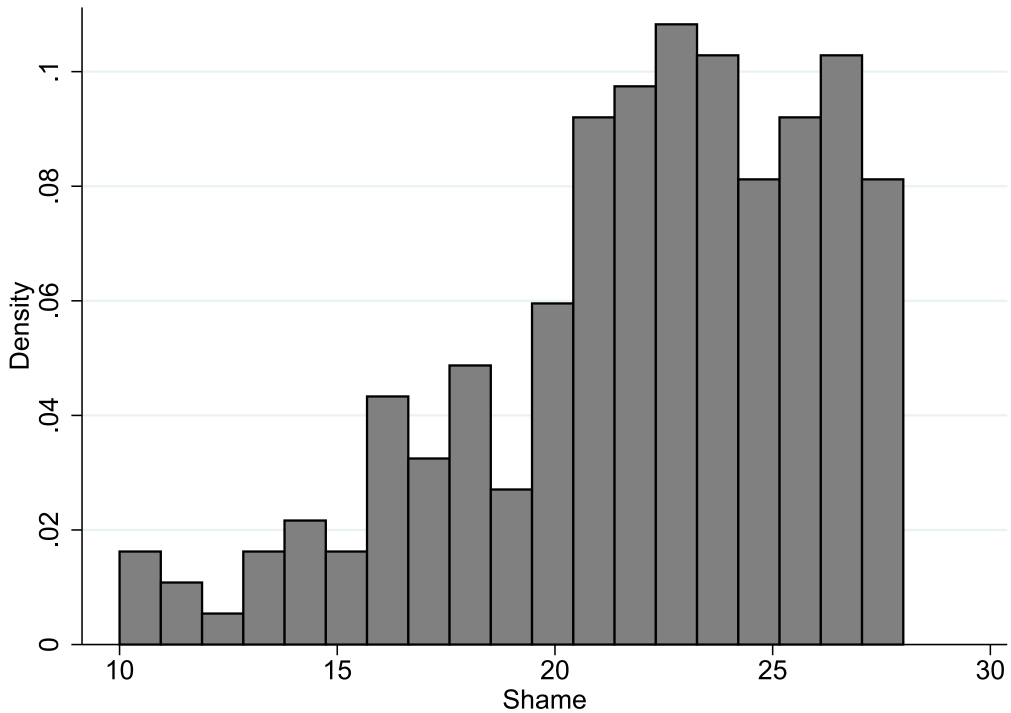


Figure C.15: Histogram of Shame Score

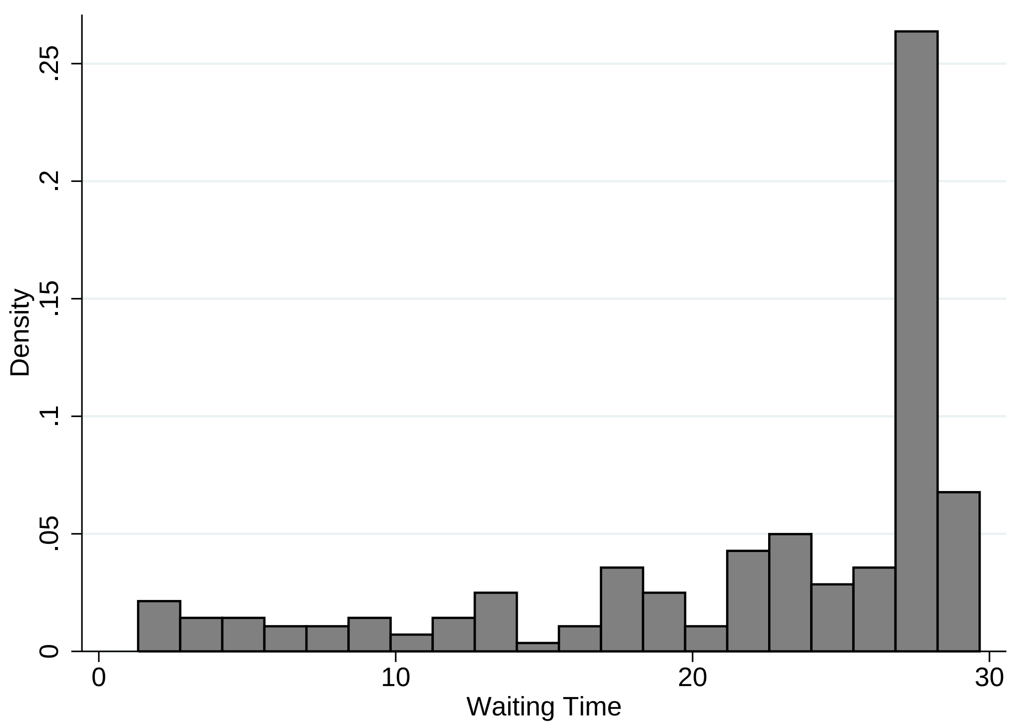


Figure C.16: Histogram of Waiting Times