THE ROLE OF STRATEGIC BELIEFS IN UNDERSTANDING STRATEGIC, PRO-SOCIAL AND SOCIALLY COMPLEMENTARY BEHAVIOR

A Dissertation

by

ADA KOVALIUKAITE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Catherine Eckel
Co-Chair of Committee,	Alexander L. Brown
Committee Members,	Daniel Fragiadakis
	Steve Puller
	Marco Palma
Head of Department,	Timothy Gronberg

May 2018

Major Subject: Economics

ABSTRACT

This dissertation includes three experimental studies discussing the role of strategic beliefs in understanding strategic, pro-social and socially complementary behavior.

The first essay, "Testing Stepwise Reasoning", introduces a test of the underlying assumption in the Cognitive Hierarchy model of strategic behavior: an individual anticipates facing a distribution of opponents that differ in terms of their strategic sophistication. In phase one of the experiment, subjects play games. Different participants in phase two then predict their behavior. We find that a subject often possesses beliefs consistent with the Cognitive Hierarchy model. We conclude that Cognitive Hierarchy is more than *"as if"* model of behavior – it illustrates, how agents *actually think* in a strategic environment.

The second essay, "Measuring Trust: A Reinvestigation", investigates two ways of measuring trust: a survey and an incentivized game. Prior literature established that the two measures are not correlated. However, a common study in this body of research employed a modified version of the original trust game, thus, potentially changing the motives and *strategic beliefs* that drive trusting behavior. We conduct a replication and a reinvestigation of this study and find that the two measures are correlated, when the original trust game is used. We suggest that trust is a single construct measured by both the survey and the incentivized game.

The final essay, "Norm Misperceptions and Social Network Structure" presents a test of a potential mechanism behind social norm misperceptions. In a laboratory, we exogenously manipulate the structure of the imposed social networks. We observe a positive bias in the socially complementary consumption and perceptions about the average consumption of others in the treatment condition, where the correlation between the induced preferences and the number of individual network links is positive. However, a positive bias also exists in the control condition, where the number of social links is fixed for all subjects. We discuss potential sources of the bias in the control condition and suggest a direction for future work. An improved design may act as a platform for testing methods to battle social norm misperceptions and their negative consequences.

DEDICATION

To my parents, Jurate Kovaliukiene and Vilmantas Kovaliukas:

You have done everything to help me succeed. This would not have been possible without your love, support, sacrifice and understanding. I hope this achievement stands as evidence that it is one of my deepest desires to make you proud. Please know I will forever be grateful for the relationship that we have, and I can only hope that I could one day be a parent that you are to me to the kids of my own.

To my late grandmother, Danute Kovarskiene:

I cannot express my gratitude for the love, the caring, and the worrying she gave me. It breaks my heart to know that she did not get to see me finish my studies, and she did not get to see me come back home once again. I will always miss her humor, her light heartedness and lively nature. I hope her bright memory will inspire my beloved grandfather Henrikas Kovarskas and her daughters, Jurate Kovaliukiene and Rasa Skrebiene, to stay strong.

ACKNOWLEDGEMENTS

This dissertation is a result of a tremendous amount of effort, not only my own, but also of the people surrounding me. Here I want to acknowledge the contribution of all of you and promise to do everything in my power to provide the help you need in your future endeavors.

Firstly, I would like to thank my committee members Dr. Daniel Fragiadakis, Dr. Catherine Eckel, Dr. Alex Brown, Dr. Steve Puller and Dr. Marco Palma for their guidance and support throughout the course of this research. I have been lucky enough to collaborate with you on joint research projects that started my path as a scholar and/or observe you as great advisers to my fellow doctorate students. I have learned very different, yet equally valuable lessons from each of you. These lessons have shaped my view of scholarship and defined my understanding of academic success.

I want to express my tremendous gratitude to Dr. Daniel Fragiadakis for being unbelievably generous with his precious time, for reminding me again and again of how fun the process of discovery can be, and for making me believe in myself even in the moments of struggle. Your positive attitude and the sincere desire to help have pushed me through the good and bad times in the program. I greatly appreciate your patience and respect for the millions of questions I have had, and the care you treat your research with. It has been an honor to be your first advisee and a great fortune to have you as my adviser.

To Dr. Catherine Eckel, thank you for sharing your vast experience and the honest truths of our profession as well as maintaining your faith in my ability. Your brilliance as a member of the academic community is truly inspiring. Over the years, you have touched many people's lives, and I am no exception. With your encouragement, I push forward, and I could not be more grateful for your support.

To Dr. Alex Brown, thank you for entertaining my pursuit of the new and the risky research paths. You have been courageous enough to embark on such a journey with me, and I am grateful for the clever insights that you shared with me along the way. I am also deeply grateful for sparking my interest in game theory, as an undergraduate student, and for the kindness you showed me only to lift my spirit. Without you, I may not have gotten where I am today.

A special thanks to Dr. Steve Puller for taking me under his wing in the very beginning of the program. You are a great inspiration as an academic for your dedication to your work. Without doubt, you are still the greatest instructor I have ever had in the eight years of my experience as a university student. I will always try to live up to the standard you set in my mind.

To Dr. Marco Palma, thank you so much for your enthusiasm, friendly nature, attention and helpful insights in the presentations of my work. I have always admired your level of engagement with our department and your students. They are absolutely lucky to have you.

Throughout this experience, many individuals have had a great influence on me. I want to extend my gratitude to the late Dr. John Van Huyck, who was the first to get me intrigued with the research in experimental economics. He recognized the excitement in me, and his recognition of my enthusiasm inspired me to pursue more. I want to thank Dr. Mark Hoekstra for bringing me into the graduate program at Texas A&M and for all the opportunities that I have had as a result. To my co-author, Dr. David Rojo Arjona, thank you for your expertise, excitement with research and support in helping me get my job market paper out of the door. To Dr. Jason Lindo and Dr. Andrew Barr, thank you for guiding me through the stressful experience of the job market. To Dr. Timothy Gronberg, Dr. Dennis Jensen and Dr. Silvana Krasteva, thank you for the awards, acknowledgements and funding that I received to support my work as a doctorate

student. To the entire faculty as well as the administrative staff at the Department of Economics at Texas A&M University, thank you for your kind words, for your feedback, your time, your faith in me, my effort and my work, and for everything that you have done to make this journey easier for me.

I want to thank the Economic Research Laboratory group for creating a collaborative research environment. I would not have run a single experiment and would not have presented a single decent presentation without your help. A special thanks to Cheryl Mitchel, who has made all of our lives so much more organized. To our former students, thank you for passing us your knowledge and sharing your experiences. To our current students and, especially, my cohort members, thank you for sharing this journey with me – it is manageable merely because we are in the same boat. Thanks to Abigail Peralta and Jose Gabriel Castillo for sharing similar attitudes towards work and life, to Noah Bacine for the tough conversations, to Billur Aksoy for being a role model of determination, to Haley Harwell for coaching me and Billur through publishing our first paper, to Sam Bondurant for the game nights. Special thanks to Jon Scott for surviving the program together – I may not have made it through my first year without your help. I am grateful to your family for having faith in us both.

I would also like to express my gratitude to the teachers and students of Vilnius Lyceum. I grew up with you. You nourished my passion for learning and developing as a human being. Without the start that you gave me in my life, I would not have taken the steps that got me where I am today. You rewarded me with a group of closest friends that support me until this day in my pursuits. To all of my friends back home – thank you.

To all of those I can call friends on this side of the ocean – thank you too. Your friendship supported me effortlessly, without you even knowing about it.

I am especially grateful to Héctor Iván Rodríguez González for helping me find home away from home, for resonating with my view of the world and making me feel like anything is possible. Thank you for your unconditional support and for your patience. The joy you brought into my life fueled me with determination for the finish line.

I must thank all of my family: my amazing grandparents, my ants and uncles, my cousins; and the friends of my parents, who were there for me, emotionally and financially, welcoming me back home each time I return and supporting me, when I leave.

Above all, I want to thank my parents, Jurate Kovaliukiene and Vilmantas Kovaliukas, who raised me, loved me, put me through school, and supported me in every step of the way with complete confidence and trust. I could never repay the debt that I owe you. However, I can fight every day to make you proud. Please know that I am immeasurably proud of the two of you. You are the greatest parents anyone could ever wish for.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professors Daniel Fragiadakis, Catherine Eckel, Alex Brown, and Steve Puller of the Department of Economics and Professor Marco Palma of the Department of Agricultural Economics.

The study in Chapter I was conducted in collaboration with Dr. Daniel Fragiadakis and Dr. David Rojo Arjona from the Department of Economics at Leicester University. The project in Chapter II is a collaboration with Billur Aksoy, Dr. Haley Harwell, and Dr. Catherine Eckel. Dr. David I. Laibson kindly provided the materials, data and the documentation of analysis from the original project replicated and reinvestigated in Chapter II. Chapter III is joint work with Dr. Alex Brown.

Graduate study was supported by a fellowship from Texas A&M University, E. Ralf Daniel Scholarship awarded by the Private Enterprise Research Center, and the research funds awarded by the Department of Economics as well as funds from the Professors Daniel Fragiadakis, Catherine Eckel, Alex Brown, and Steve Puller.

NOMENCLATURE

BA	Box Arrangement
BR	Bomb Risk
C	classified
CH	Cognitive Hierarchy model
ERL	Economic Research Laboratory
GLSS	Glaeser, Laibson, Sheinkman and Sutter (2000)
LK	Level *k* model
MR	Money Request
NS	Number Selection
RP	Relative Performance
TAMU	Texas A&M University
TG	Trust Game
U	unclassified
QSR	Quadratic Scoring Rule

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. GENERAL INTRODUCTION

This dissertation includes three essays investigating the role that strategic beliefs play in determining individual behavior. Previous work has provided evidence that subjective strategic beliefs are meaningful in understanding individual behavior in situations of strategic uncertainty. The work I present here expands this body of research by discussing the significance of strategic beliefs in the contexts of strategic, pro-social and socially complementary behavior.

In the first essay (Section 2), we study the strategic beliefs an individual possesses in a simple strategic environment. One of the most prominent belief-based game theory models of behavior, that of Cognitive Hierarchy (Camerer et al., 2004), characterizes an individual as believing her opponents engage in heterogeneous steps of strategic thinking. We are the first, to our knowledge, to directly test this assumption. In Phase 1 of our experiment, subjects play games. Their behavior is then predicted by a group of different participants in Phase 2. A subject appears to believe in a handful of non-strategic individuals who play a naive strategy. She also, however, expects there are players who play the best-response to this strategy as well as types who best respond to these strategic individuals and so on. We thus find that Cognitive Hierarchy's beliefs assumptions clearly pass our test. We further suggest that the assumed stepwise reasoning illustrates how individuals *actually think* in a simple strategic environment, thus, providing strong foundation for the belief-based strategic models of behavior.

In the second essay (Section 3), we discuss how malleable strategic beliefs and other motives guiding trusting behavior may be easily affected by the experimental design parameters, which can lead to mismeasuring trust. In particular, we reinvestigate the question first posed by Glaeser, Laibson, Scheinkman and Soutter (2000, GLSS hereafter): do survey measures about trust predict actual trusting behavior? This important study established that the behavior in an incentivized trust game is not correlated with the responses to the most widely used survey measures of trust. However, GLSS employed a modified version of the original Berg et al. (1995) trust game, which may have changed both the strategic beliefs and other motives that give rise to trusting behavior in this game. We conduct a replication and a reinvestigation of GLSS. In the replication, we use the GLSS protocol and we reproduce their results. In the reinvestigation, we introduce one major change: we replace the GLSS version of the investment game with the standard Berg et al. (1995) game. The standard game endows both players, while the modified version endows only the first mover. After endowing both movers in the reinvestigation experiment, we find a statistically significant correlation between the two measures of trust. We conclude that it is important to take care in designing experiments measuring complex economic concepts, such as trust, especially if the measured behavior is heavily based upon malleable strategic beliefs. In doing so we discover that trust is a single construct, whether measured by the survey questions or by an incentivized trust game.

In the final essay of my dissertation (Section 4), we investigate the relationship between the accurate strategic belief formation and optimal behavior. Literature in economics and sociology suggests that in some social contexts individuals may possess biased beliefs regarding the behavior of the population. Such norm misperceptions are, in turn, associated with excessive engagement into socially undesirable behavior, such as alcohol abuse or smoking. Jackson (2016) proposes that, in the context of socially complementary behavior, norm misperceptions may arise due to the social network structures we interact in. In this paper, we present a laboratory experiment, where we exogenously manipulate the structure of the networks our participants interact in. We demonstrate that subjects exhibit overconsumption of an abstract, socially-complementary good and upwards-biased beliefs regarding the consumption of others when the correlation between the induced preferences and the number of individual network links is positive. However, arguably smaller bias also exists in the control condition, where the number of social links is fixed for all subjects. We discuss potential sources of the bias in the control condition and outline design adjustments to eliminate them. The improved design may act as a platform for testing methods to battle social norm misperceptions and the associated overengagement into socially complementary activities.

Together, my three essays investigate the importance of strategic beliefs in understanding strategic, pro-social and socially complementary behavior. Through direct belief elicitation and theoretical reasoning in an experimental setting, we learn that understanding strategic beliefs and their formation is informative to economic theory, useful in refining methods to measure complex economic concepts and essential in planning policy interventions.

2. TESTING COGNITIVE HIERARCHY ASSUMPTIONS

2.1 Introduction

Camerer et al. [2004] develop a Cognitive Hierarchy (CH) model of behavior in one-shot games and use it to characterize non-equilibrium play typically observed in laboratory experiments.[1] CH has since also been used to estimate behavior in the field.[2] Though used to model aggregate data, CH makes individual-level assumptions. Specifically, it describes a player as a Step k thinker who best-responds to a belief that others are Step *0* through Step $k - 1$ thinkers, where Step *0* thinkers play naive strategies.[3] To our knowledge, direct evidence that an individual holds such beliefs is nonexistent in the sense that we are unaware of a study that tests whether an individual anticipates thinkers of several steps when asked to forecast the play of a population. Such a test is important, however, if CH is intended to be a descriptive model. In this paper, we conduct such a test using a lab experiment.

In our Actions treatment, subjects play a series of two-person Number Selection (NS) games with no feedback. In NS games, players simultaneously select integers, say, between 1 and 14, inclusive. If a player selects an integer, *i*, she earns *i* points.[4] Furthermore, a player earns 100 bonus points her number is exactly 3 (in some games, 4) less than her opponent's number and earns 35 points if her number equals her opponent's number.[5] We expect a non-strategic Step *0* thinker to play 14 in this example, since this is the best number to choose if one does not

---

1 See Camerer (2003) for a behavioral game theory overview.
2 For example, Brown et al. (2013) use CH to more accurately predict moviegoer behavior in comparison to equilibrium. Hortacu et al. (2017) estimate firms' levels of strategic sophistication using the CH model and find that larger firms engage in more complicated reasoning compared to smaller firms.
3 Stahl and Wilson 1995 also describe an individual as having diverse beliefs over opponents.
4 In both treatments, earned points are used to pay subjects with binary lotteries (see Roth and Malouf, 1979), incentivizing a subject to maximize her expected number of points, irrespective of her risk preferences.
5 Our NS games are inspired by others from Arad and Rubinstein (2012), Georganas et al. (2015) and Fragiadakis et al. (2017).

consider how his payoff is affected by his opponent's play.[6] A Step 1 player, then, who assumes her opponents are Step *0* types, is incentivized to choose 11. Step *2* players anticipate Step *1* and Step *0* opponents; thus, depending on the particular distribution of beliefs held by a Step *2* player, she may have a best-response of 11 or 8. Continuing this procedure, it can be shown that the predictions from Step *k* reasoning are confined to $S = \{14,11,8,5,2\}$. Furthermore, a Step *k* thinker who anticipates others to play some s ∈ S should also expect strategies $\{s' | s' > s \text{ } and \text{ } s' \in S\}$ to be played. This provides us with a clean way to test for Step *k* beliefs in our next treatment.

In our Beliefs treatment, each participant states her beliefs regarding the play of a set of 20 Actions subjects in each of the NS games. To state her beliefs for a particular NS game, a Beliefs participant constructs a 20-box histogram over a horizontal axis that displays each of the game's pure strategies. For constructing a given histogram, *h*, a subject earns *p* points, where *p* is the number of boxes that would be overlapping were we to place *h* on top of the histogram of actual behavior by the 20 Actions subjects. This renders our mechanism incentive compatible: if a Beliefs subject thinks *s* subjects chose the number n, she is incentivized to place *s* boxes above *n* when constructing her histogram. We also record the order in which Beliefs participants arrange their boxes in their histograms.[7] Despite having no effect on payoffs, we expect a Beliefs participant to place boxes on lower thinking steps earlier.

---

[6] An alternative rule-of-thumb for a Step *0* thinker is to select an integer uniformly at random. Either of these Step *0* specifications give rise to a best-response of guessing 11 in this example, thus, both rules lead Step *1* thinkers to the same number.

[7] The order in which boxes are arranged does not affect payoffs and the instructions make no mention of the order in which boxes are arranged. We view this minimally invasive "belief-tracking" procedure as relating to studies using "eye-tracking" to record where subjects direct their attention (see Wang et al., 2010) as well as studies where subjects must actively "open" boxes to observe payoffs from certain strategy profiles (see Costa-Gomes et al., 2001).

We find strong support for Step k beliefs: for 48% of our Beliefs participants, conditional on anticipating k steps of reasoning, k' steps of reasoning are also expected for all $k' \leq k$. Illustrating this using our previous NS game example, if a Beliefs participant expects sufficiently many Actions subjects select the number 5, she also believes sufficiently many select 14, 11 and 8. In terms of our "belief-tracking" data, we find that beliefs indeed follow stepwise reasoning: when a box is placed on a certain step of thinking, the subsequent box is likely to be placed on that same step or one step higher.

The remainder of the paper is organized as follows: Section 2.2 presents the experiment, Section 2.3 discusses the results, Section 2.4 mentions some related literature and Section 2.5 concludes.

2.2 The Experiment

2.2.1 Number Selection (NS) Games

We carefully design Number Selection (NS) games for Actions participants to play. Before describing NS games in detail, we discuss a design challenge to overcome if we seek to cleanly identify Step k thinking in the Beliefs treatment. Suppose Actions subjects were to play games where idiosyncratic Step k reasoning would generate a variety of best responses. If a Beliefs subject thinks Actions participants are idiosyncratic Step k thinkers, she will expect Actions data to be thinly scattered. This would make it difficult for us to distinguish such a highly strategic Beliefs participant from another who randomly constructs a histogram non-strategically. Using a game with a dominant strategy would solve this issue but would introduce the "opposite" problem since it would merge all Step $k \geq 1$ actions, not allowing us to distinguish a Beliefs participant who only anticipates Step *0* and Step *1* Actions participants from another who believes Actions subjects engage in additional thinking steps.

We thus design NS games so that Step *k* predictions are neither too diluted nor too concentrated. They satisfy the following property: a Step *k* thinker – for any risk attitudes and any beliefs of how Step *0* through $k - 1$ types are distributed – should only select actions consistent with the Level *k* model from Stahl and Wilson (1994) and Nagel (1995) (see Observation 4). [8]

The Level k action in a game is calculated by taking the best-response to the Level $k - 1$ action; Level *0* players are nonstrategic like Step *0* thinkers. This ensures that if a Beliefs subject thinks that Actions participants are all highly heterogeneous Step *k* reasoners, she should only expect Actions subjects to select Level *k* actions. The Level *0* through *4* actions in the particular NS games we use are all distinct. [9]

In a generic NS game, g, a player *i* and her opponent simultaneously select integers, n_i and n_{-i}, respectively from a common range $R_g = \{1,2, \ldots, UB_g\}$, where UB_g is the game's upper bound. Player *i* earns n_i points automatically for selecting n_i. If n_i is exactly D_g less than n_{-i}, where D_g is g's commonly known undercutting distance, then *i* earns $B_g > UB_g \times D_g$ additional points.[10] If $n_i = n_{-i}$, then player *i* earns $b_g \in (UB_g - 1, B_g - D_g)$ additional points.[11] This payoff function is shown in Equation 1.

---

[8] This assumes a Level *0* and Step *0* action of choosing the upper bound in our NS games, which we justify in the ensuing paragraphs.

[9] While we could have expanded our games such that the Level *0* through *k* actions in each NS game are all distinct for some $k > 4$, we focus on Levels *0* through *4* because prior evidence shows that play at higher levels drops off rather quickly. For example, Crawford and Costa-Gomes (2006]) and Fragiadakis et al. (2016) classify substantially more Level *1* and Level *2* players in comparison to Level *3*.

[10] The restriction that $B_g > UB_g \times D_g$ is needed for Observation 2.

[11] The restrictions that $b_g < B_g - D_g$ and $b_g > UB_g - 1$ are needed for Observations 1 and 3.

$$\pi_i{}^g(\mathrm{n}_i, \mathrm{n}_{-i}) = \mathrm{n}_i + \begin{cases} B_g \; if \; n_i = \; n_{-i} - D_g \\ b_g \; if \; n_i = \; n_{-i} \end{cases} \qquad (1)$$

The Level k model provides clear behavioral predictions in NS games. We assign UB_g as the Level *0* (and Step *0*) action,[12] giving rise to the Level k actions stated in Observation 1.

Observation 1. In a Number Selection game, g, the Level k strategy is max $\{UB_g - k{\times}D_g, mod(UB_g, D_g)\}$, where $mod(x, y)$ is the remainder from $x \div y$.

Proof. The Observation states that when undercutting a Level $k-1$ player is not possible, the Level k action coincides with the Level $k-1$ strategy. Thus, assume that an opponent selects $n_{-i} \leq D_g$ (so that undercutting is not possible). Then, $b_g > UB_g - 1$ implies $\pi_i{}^g(n_{-i}, n_{-i}) = n_{-i} + b_g > n_{-i} + UB_g - 1 \geq UB_g = \pi_i{}^g(UB_g, n_{-i})$. The Observation also states that when undercutting a Level $k-1$ player is possible, the Level k best response is to undercut. Assume an opponent selects $n_{-i} > D_g$ (so that undercutting is possible). Given that $b_g < B_g - D_g$, we have $\pi_i{}^g(n_{-i}, n_{-i}) = n_{-i} + b_g < n_{-i} + B_g - D_g = \pi_i{}^g(n_{-i} - D_g, n_{-i})$.

While the Level k predictions in Observation 1 assume an upper bound Level *0*, there may be alternative strategies that naive subjects would implement. For instance, a player with absolutely no understanding of the NS games may play uniformly at random. This Level *0* specification is actually quite common. We design our experiment such that this alternative Level *0* specification does not alter the Observation 1 predictions for $k \geq 1$, minimizing concerns about the explanatory power of Level k being driven by idiosyncratic Level *0* specifications (see

---

[12] Arad and Rubinstein (2012), who have a similar game, also designate the upper bound in their games as Level *0* since it is the action that maximizes a player's payoff if he does not form beliefs about his opponent and acknowledges only that his payoff function grants him n points for selecting n.

Hargreaves Heap et al., 2014).

Observation 2. For each Number Selection game, *g*, selecting $UB_g - D_g$ is the unique action that maximizes one's expected points against an opponent who plays uniformly at random.

Proof. Against a uniform random opponent, player *i*'s points in game *g* are:

$$\Pi_i{}^g(\mathrm{n}_i) = \frac{1}{UB_g}\sum_{j=1}^{UB_g} \pi_i{}^g(\mathrm{n}_i, j) = \mathrm{n}_i + \frac{b_g}{UB_g} + \begin{cases} \frac{B_g}{UB_g}\ if\ \ \mathrm{n}_i \le UB_g - D_g \\ 0\ if\ \ \mathrm{n}_i > UB_g - D_g \end{cases} \tag{2}$$

From Equation 2, we see that UB_g and $UB_g - D_g$ are the only local maxima. Given that $B_g > UB_g \times D_g$, we have $\frac{B_g}{UB_g} > D_g$, which implies that

$$\left(UB_g - D_g + b_g/UB_g\right) + \frac{B_g}{UB_g} > \left(UB_g - D_g + b_g/UB_g\right) + D_g$$

Therefore, $\Pi_i{}^g\left(\mathrm{UB}_g - \mathrm{D}_g\right) > \Pi_i{}^g\left(\mathrm{UB}_g\right)$, rendering the Level 1 prediction (and hence all Level $k \geq 1$ actions) from Observation 1 unchanged.

In Observation 1, we reason that if a player cannot undercut his opponent, her best response is to match the other player. This proves Observation 3.

Observation 3. A Number Selection game with undercutting distance D_g has D_g pure strategy Nash equilibria: $\{(1,1),(2,2),\ldots,(D_g,D_g)\}$.

Before moving to the Results section, we formally define the Step *k* model used in this paper, which is similar to those from Stahl and Wilson (1995) and Camerer et al. (2004). Step *0* and Step *1* players are identical to Level *0* and Level *1* types, respectively. For $k > 1$, a Step *k* thinker anticipates Step *0* through Step *1* players (with any proportions). Observation 4 outlines the behavioral predictions of Step k thinkers in NS games.

Observation 4. Suppose a player i (in a Number Selection game) believes the chance of facing a Level k action is $a_k \in [0,1]$ for all $k \geq 0$. If $\sum_{h=0}^{\infty} a_h = 1$, her best response (given her beliefs) is some Level k action.

Proof. Suppose player i selects some non-Level k action, n_i, in an NS game, g. Her payoff will be n_i since there will be no way of receiving the b_g or B_g bonuses. As a result, deviating to $UB_g > n_i$ is profitable.

2.2.2 Box Arrangement (BA) Tasks

We create Box Arrangement (BA) tasks to elicit beliefs. The objective in a BA task t_g that corresponds to a game *g* (with finite strategy set *S*) is to distribute N_g boxes across a horizontal axis made up of the pure strategies in *g*. Let *h* denote the histogram generated, where h_s is the number of boxes that placed on pure strategy *s*. Let *l* denote the histogram of actual behavior by some population of N_g individuals in game *g*, where l_s is the number of times *s* was chosen. The points earned in t_g are equal to the number of boxes that are overlapping when *h* and *l* are superimposed. Equation 3 gives the payoffs function mathematically.

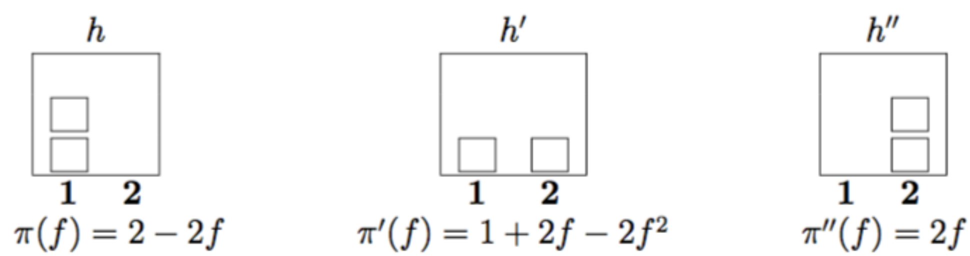
$$\pi\,(h, l) = \sum_{j=1}^{|S|} \min\,\{h_j, l_j\} \tag{3}$$

Thus, if *h* perfectly coincides with *l*, N_g points are earned. If *h* does not coincide with *l* whatsoever, 0 points are earned. Suppose a risk neutral individual *i* has a belief system *α* for a NS game, g.[13] In other words, she believes the chance a randomly chosen opponent will select

[13] We control for risk attitudes in our experiment, as described in Section 2.3.

$s \in R_g$ is $a_s \in [0,1]$. We thus assume that $\sum_{h=1}^{UB_g} \alpha_h = 1$. The payoffs in Equation 3 incentivize boxes to be arranged in a manner that approximates the beliefs system α. We illustrate this in Figure 1 with an example where $R_g = \{1, 2\}$ and $N_g = 2$. An individual can create h, h' or h''. Let $f \in [0,1]$ denote the fraction of the population that is believed will select 2. It can be shown that the expected payoffs of building histograms h, h' and h''. (as functions of f) are π, π' and π'', respectively. Furthermore, it can be shown that (i) for $f \leq 1 - \frac{\sqrt{2}}{2}$, $\pi(f) \geq \max\{\pi'(f), \pi''(f)\}$. Thus, for sufficiently small f, one is best off by placing both boxes on 1. Symmetrically, for sufficiently large f, one's payoff-maximizing histogram involves both boxes placed on 2; in other words, (ii) $\pi''(f) \geq \max\{\pi'(f), \pi(f)\}$ for $f \geq \frac{\sqrt{2}}{2}$. If f takes on intermediate values, it is optimal to place a box on 1 and a box on 2; it can be shown that (iii) for $f \in \left[1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right]$, we have $\pi'(f) \geq \max\{\pi''(f), \pi(f)\}$.

Figure 1: Histograms and payoff functions for $R_g = \{1, 2\}$ and $N_g = 2$.



2.2.3 Treatments and Procedures

Experimental subjects are undergraduates and graduate students, recruited using ORSEE (Greiner et al., 2015). Participants interact via a network of computers linked by z-Tree (Fischbacher, 2007) at the Economics Research Laboratory in TAMU's Department of Economics. Two 20-subject sessions, each lasting approximately 1 hour, make up the Actions treatment. The 81 participants in the Beliefs treatment are spread across 5 sessions of 14, 13, 18, 18 and 18 subjects, each taking roughly 2 hours. Average earnings are \$28.39 and \$54.07 in the Actions and Beliefs treatments, respectively, including a \$5.00 show-up payment.

Subjects in the Actions treatment play the 11 Number Selection games shown in Table 1, where each game has B_g = 100 and b_g = 35. The games we select are determined using a list of design criteria. First, each game's lower bound is 1 and upper bound is no more than 32 due to spatial constraints on subjects' computer screens. Second, we impose that $D_g \in \{3, 4\}$ so that the Level *k* model only captures a fraction of the strategy space. Third, we make it so that $\{UB_g - k \times D_g\} \cap \{1, D\} = \emptyset$ for all k,which prevents the Nash equilibrium reached via the Level k model from coinciding with the "lower bound equilibrium" of (1,1) or the "efficient equilibrium" of (D_g,D_g). Subjects may play "1" as a rule of thumb or as a focal final result of repeated (but not iteratively performed) undercutting. Participants who are "true" equilibrium players that do NOT arrive at equilibrium via converged Level *k* reasoning may be expected to coordinate on (D_g,D_g), the most efficient equilibrium. Lastly, $UB_g \geq 14$ ensures that Levels *0* through *4* are all distinct (though Level 4 may coincide with an equilibrium).

Table 1: The 11 Number Selection games used in the experiment

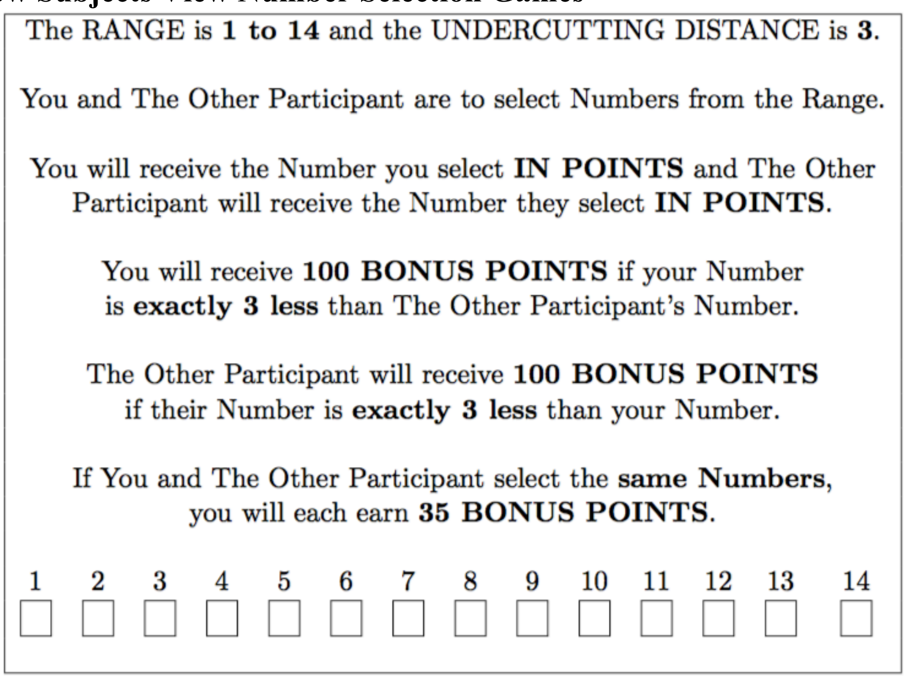
Game Number	1	2	3	4	5	6	7	8	9	10	11
UB_g	14	17	20	23	26	29	32	18	19	22	23
D_g	3							4			

Actions Participants in a session are paired randomly and anonymously at the beginning of their session; matching is fixed for all 11 games. NS games are presented to participants as in Figure 2. Each subject in the Actions treatment views a NS game from the same perspective: a subject is addressed as "You" and her opponent is referred to as "The Other Participant". In the Beliefs treatment, participants are shown the same 11 NS games from Table 1, presented as in

Figure 2, except that the two players are referred to as "Jack" and "Jill".[14]

For a game, g, a Beliefs subject performs a BA task, t_g. When performing a BA task, a subject's screen initially shows a large, empty, rectangular area that has the game's range of numbers, $R_g = \{1, 2, \ldots, UB_g\}$, strung along its lower horizontal edge. A green, upwards-pointing, arrow button rests underneath each number in R_g. Clicking the green arrow button under a number $n \in R_g$ adds a blue box above n. When one or more blue boxes sit above n, a red, downwards-pointing, arrow button shows below n's green arrow button. Clicking n's corresponding red arrow button removes the top-most blue box that sits above n. A subject can click green and red arrow buttons without any restrictions, allowing her to freely build (and revise) her histogram until she is ready to submit it. Between the game's range and the green arrow buttons, a counter tracks how many boxes sit above each number in R_g.

Figure 2: How Subjects View Number Selection Games



[14] Importantly, Beliefs participants are not shown any decisions made by Actions subjects.

Subjects click the arrow buttons to allocate their 20 blue boxes across the strategies in R_g to express how they believe the 20 participants from a previously run Actions session made their choices in g. For instance, if a Beliefs participant believes that two Actions subjects chose $1 \in R_g$, she would place two blue boxes above the number "1". Figure 3 shows the example Box Arrangement Task (generated at random) that we gave subjects in the Beliefs treatment instructions. To explain the incentives in Equation 3, we also provide subjects with a Dot Arrangement of hypothetical play by 20 Actions subjects. The Dot Arrangement is printed on standard white paper while the example Box Arrangement from Figure 3 is printed on a plastic transparency. This allows subjects to overlay the two arrangements to generate an image that looks like that shown in Figure 4. Subjects are then told that they earn p points for having p boxes that overlap with dots. While the points earned do not depend on the order in which boxes are arranged, we record the order in which Beliefs participants arrange the boxes that make up their final histograms.

Figure 3: A histogram for a fictitious Box Arrangement Task

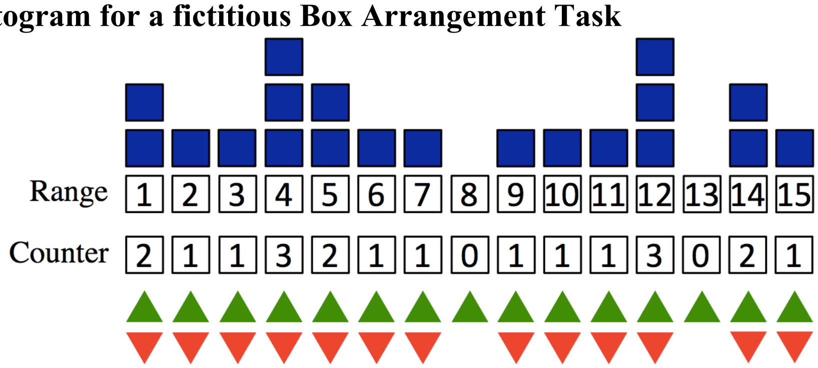
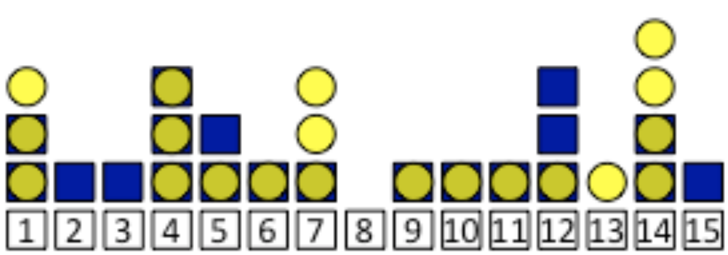


Figure 4: **A fictitious Dot Arrangement overlaid with the histogram from Figure 3**



In both treatments, instructions are read out loud and subject comprehension is reinforced using on-screen understandings tests. A subject cannot proceed past a question until it is answered correctly.[15] The incentivized portion of each treatment is initiated only once all subjects complete the treatment's understandings test. Subjects receive no feedback whatsoever between NS games and BA tasks until all experimental decisions have been made.

To mitigate concerns about risk attitudes influencing decision-making, the point payoffs in each NS game and BA task are converted to money (at the end of the experiment) using separate and independently run binary lotteries (Roth and Malouf, 1979). If a subject earns p points in a NS game, the corresponding lottery pays \$5 with probability $p/150$ and \$1 with probability $1-p/150$.[16] If a subject earns p points in a BA task, she earns \$5 with probability $p/20$ and \$1 otherwise.

2.2.3.1 Relative Performance (RP) Questions

After a Beliefs subject i completes all 11 BA tasks, she performs 11 corresponding Relative Performance (RP) questions. For the RP question q_g corresponding to BA task t_g and NS game g, subject i is shown g as well as the histogram she constructed in t_g. Participants are not shown any histograms that were made by any other Beliefs subjects. Subject i is informed of the number of subjects in her lab session and is asked to estimate how many participants in her session she believes earned strictly more points than she did in t_g. For q_g, subject i earns \$5 for a correct answer and \$1 otherwise. The RP questions are intended to estimate subjects' levels of confidence in the histograms created in the BA tasks.

---

[15] Incorrect answers are met with a prompt asking the subject to try again. If a subject is stuck, she can quietly ask for individual assistance from the experimenter.

[16] There are two main reasons for using binary lotteries in NS games. First, doing so gives rise to Observation 2. Second, a linear exchange rate of points to dollars is problematic for subject payments since a participant may earn over one hundred times as many points as another in a NS game.

2.2.3.2 Bomb Risk (BR) Decisions

After subjects in the Beliefs treatment perform their RP questions, they make a Bomb Risk (BR) decision (adapted from Crosetto and Filippin, 2013) as a quick measure of their risk attitudes. The BR decision is very straightforward. There are 100 treasures chests, one of which contains a bomb. The subject chooses, m, the number of (randomly picked) chests it would like the computer to open. If the bomb is in an opened chest, the subject earns nothing (which occurs with a $m/100$ chance). Otherwise, the bomb-filled chest is not opened and the subject earns $m/10$ dollars.

2.3 Results

2.3.1 Aggregate Behavior

2.3.1.1 Actions Treatment

Before analyzing data from the Beliefs treatment, we briefly summarize Actions behavior. Observation 4 states that, for a wide range of beliefs, individuals' best responses in Number Selection (NS) games coincide with the Level k actions. We thus expect a substantial proportion of behavior in NS games to fall on Level k actions; our data confirm this (see Result 1 and Table 2).

Table 2: Level *k* action frequencies in Actions treatment

Value of k	0	1	2	3	4	total
# of choices	60	129	89	28	4	310
% of choices	13.6	29.3	20.2	6.4	0.9	70.4

Result 1. In the Actions treatment, 70.4% of decisions coincide with the Level 0 through 4 actions; this high frequency is expected given Observation 4.

2.3.1.2 Beliefs Treatment

Roughly thirty percent of Actions behavior did not overlap with Levels *0* through *4*

(Result 1). This provides a rough estimate on the proportion of non-belief-based decision-making in the population from which we draw our experimental subjects. [17]We thus expect a baseline rate of roughly thirty percent of Beliefs data to not overlap with Levels *0* through *4*. There is another source of noise in the Beliefs data, however. Suppose that a participant *i* believes that opponents guess uniformly at random. If *i* participates in the Actions treatment, she will guess Level *1*, the most commonly selected Level *k* strategy in the Actions treatment. In the Beliefs treatment, however, she will spread out her boxes. Taken altogether, we expect Beliefs data to be more diluted than Actions behavior. Table 3 and Result 2) confirm this hypothesis.

Table 3: Level *k* action frequencies in Beliefs treatment

Value of k	0	1	2	3	4	total
# of choices	2305	2631	1635	884	597	8052
% of choices	12.9	14.8	9.2	5.0	3.4	45.2

Result 2. In the Actions treatment, 70.4% of decisions coincide with the Level 0 through 4 actions; this high frequency is expected given Observation 4.

2.3.2 Individual Analysis

2.3.2.1 Classification of Beliefs Participants as Step *k* Thinkers

We begin investigating Step *k* thinking in the Beliefs treatment by determining the Level k action(s) for $k \in \{0,1,2,3,4\}$ that a Beliefs subject anticipates being played by Actions participants. For a subject, *i*, we compute a vector, $v_i = (v^1, \ldots, v^9)$, where (i) v^{2k+1} is the total number of boxes that *i* places on the Level *k* action and for $k \in \{0, 1, 2, 3, 4\}$ and (ii) v^{2k} is the total number of boxes that *i* places between[18] the Level $k - 1$ and Level *k* action for $k \in \{1, 2, 3,$

[17] It is possible that some of the non-Level *k* choices in the Actions treatment are best responses to alternative beliefs. Similarly, some of the Level *k* choices may arise by chance.

[18] For instance, in NS game 1, where $D1$ = 3 and $UB1$ = 14, the Level 0 and Level 1 actions are 14 and 11, respectively; thus, the numbers between the Level *0* and Level *1* actions are 13 and 12.

4}. For a participant to forecast a Level *k* action, we first require that she places significantly more boxes on the Level *k* strategy in comparison to a subject playing uniformly at random (Condition 1).

Condition 1. A necessary condition for identifying a Beliefs subject *i* as believing Actions participants play the Level *k* action (for any $k \in \{0, 1, 2, 3, 4\}$) is $v^{2k+1} \geq 16$.[19]

We then require that Level *k* actions are forecasted sufficiently more than numbers in their vicinities.[20] Specifically, we create a weighted vector, $w_i = (w^1, \dots, w^9)$, where ${w_i}^{2k+1} = {v_i}^{2k+1}/11$ for $k \in \{0, 1, 2, 3, 4\}$ and ${w_i}^{2k} = {v_i}^{2k}/26$ for $k \in \{1, 2, 3, 4\}$.These normalizations account for the fact that while there are only 11 strategies across the 11 games corresponding to a particular Level *k* action (for $k \in \{0, 1, 2, 3, 4\}$), there are 26 strategies across the 11 games between the Level $k-1$ and Level *k* actions (for $k \in \{1, 2, 3, 4\}$) .[21]

Condition 2. A necessary condition for documenting a Beliefs subject i as believing Actions participants play the Level *k* action is

- ${w_i}^1 \geq 2{w_i}^2$ if $k = 0$,
- ${w_i}^{2k+1} \geq {w_i}^{2k} + {w_i}^{2k+2}$ if $k \in \{1,2,3\}$, and
- ${w_i}^9 \geq 2{w_i}^8$ if $k = 4$.[22]

---

[19] The threshold of 16 is reached via simulations. We generate 10,000 artificial Beliefs subjects who allocate their boxes uniformly at random and find that 95% of subjects place fewer than 16 total boxes in total across the upper bounds of the 11 games.

[20] Consider, for instance, a subject who places 2 boxes on each of the 10 largest numbers in each NS game. It is not clear that this subject is a Step *k* thinker, yet he will have 22 boxes on the Level *0, 1* and *2* actions and hence meet Condition 1 for these three strategies.

[21] $D_g = 3$ for $g \leq 7$. Since there are 2 and 3 numbers between the Level $k-1$ and Level *k* actions for $k \in \{1,2,3,4\}$ in games with $D_g = 3$ and $D_g = 4$, respectively, this yields 7×2+4×3 = 26 total numbers across all histograms.

[22] We do not require that the mass placed on Level *4* is greater than the sum of (i) the mass placed between Level *3* and *4* and (ii) the mass placed between Level *4* and *5* because, in some games, there are different numbers of pure

Definition 1 states that if a subject's stated beliefs meet Conditions 1 and 2 for a Level k strategy, then she believes that strategy is played by Actions participants. For example, for subject i to believe in the Level 3 action, it is necessary that i place $b \geq 16$ boxes on the Level 3 action (totaled over all her histograms) and that $b/11 \geq s/26$, where s is the number of boxes placed between the Level 2 and Level 3 actions and between the Level 3 and Level 4 actions (totaled over all her histograms).

Definition 1. For any $k \in \{0,1,2,3,4\}$, a Beliefs participant *i* believes Actions participants play the Level *k* strategy if Conditions 1 and 2 are met for that Level *k* strategy.

Using Definition 1, we can check whether Beliefs participants build histograms that are consistent with the predictions of the CH model. In other words, we can define Step *k* thinking (Definition 2) and check for Beliefs participants who employ it (Result 3).

Definition 2. A Beliefs participant is a Step *k* thinker, for some $k \in \{1,2,3,4,5\}$, if she believes Actions participants play the Level *k* strategies (according to Definition 1) for all $k' \in \{0, \ldots, k-1\}$ and there is no larger $k \in \{2,3,4,5\}$ satisfying this property.

Result 3. We classify 40.7% of Beliefs participants as Step *k* thinkers for $k \in \{2,3,4,5\}$, supporting the assumptions in Stahl and Wilson (1995) and Camerer et al. (2004).

Table 4: Classification of types in Beliefs treatment

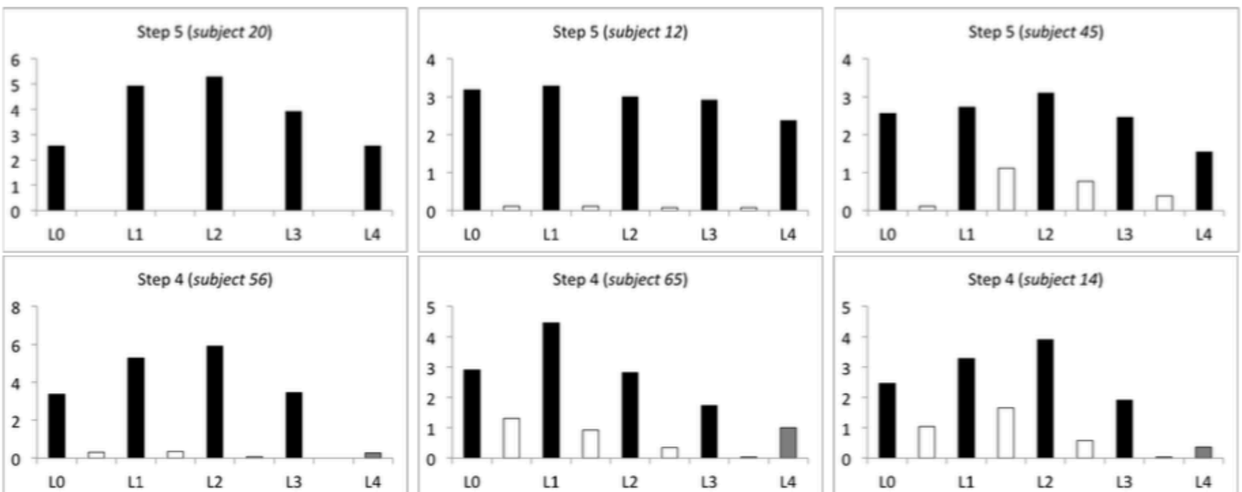
Value of *k*	1	2	3	4	5	total
# of participants	6	15	12	3	3	39
% of participants	7.4	18.5	14.8	3.7	0.9	48.1

To see Step *k* thinking more strikingly, we plot Step *k* thinkers' w_i vectors. Figure 5 does

strategies in the regions described by (i) and (ii). For example, in the first NS game, the actions 3 and 4 form the region described by (i) while the number 1 is the only action in the region described by (ii).

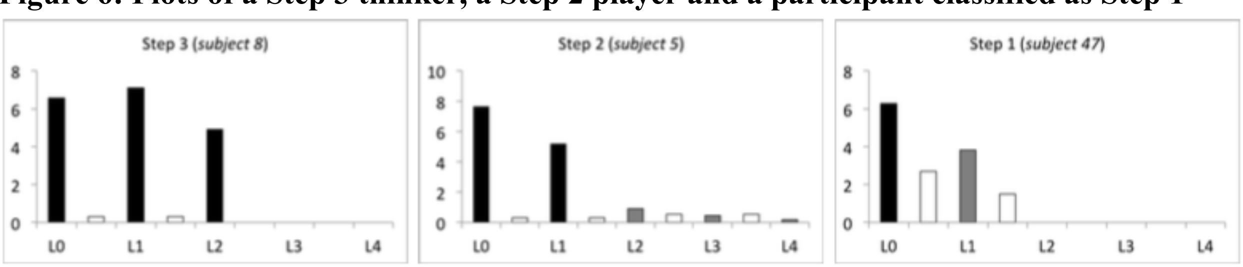
this for the three Step 5 thinkers and three Step 4 players classified in Table 4. Notice that a panel has nine categories on the horizontal axis, representing the nine coordinates of w_i. In a panel constructed by a subject, i, the black bars denote the actions anticipated by i. For example, subject 1 is a Step *5* player, and hence, Level *0* (L0) through Level *4* (L4) are black. Bars on Level k actions that are not anticipated are in gray, such as the L4 bar in subject 5's panel. White bars represent the mass between Level k actions.

Figure 5: The three Step 5 thinkers (top) and three Step 4 players (bottom) classified in the Beliefs treatment.



Examining the panels in Figure 5, we see that the black bars meet Condition 1: each black bar exceeds 16/11 ≈ 1.45. (We see that the gray bars do not meet this condition.) We can also readily see that Condition 2 is met based on how disproportionately taller the black bars are than the white bars. Figure 6 plots the w_i vectors for a Step 3, 2 and 1 thinker.

Figure 6: Plots of a Step 3 thinker, a Step 2 player and a participant classified as Step 1



2.3.2.2 Analyzing Unclassified Participants

In addition to our 39 Classified (C) participants, we have 42 Unclassified (U) subjects in the Beliefs treatment. Given that we find a non-negligible number of U participants, we would like to devote some attention to understanding their behavior. We begin by asking whether they anticipate deterministic types that we do not understand with existing models. Several findings suggest that this is unlikely. If we partition the columns built in the Box Arrangement (BA) tasks by height, we find that the taller a column is, the more likely it is built on a Level *k* action. This is shown in Figure 7. The x-axis indicates column height; the y-axis shows the percentage of columns (of that height) that are constructed on a Level k action. In other words, the blue dot corresponding to 8 along the x-axis has a height of just over 80%, meaning that roughly 80% of the columns that are exactly 8 boxes tall are built on Level *k* actions. Notice that each dot with an x-value of 10 or greater has y-value Level *k* frequency of 100% (Result 5). It thus appears unlikely that U participants are anticipating deterministic types that our existing models fail to capture.

Result 4. If a column from a histogram in the Beliefs treatment consists of 10 or more boxes, that column rests on top of a Level *k* strategy.

Figure 7: Percentage of Level *k* beliefs according to belief strength

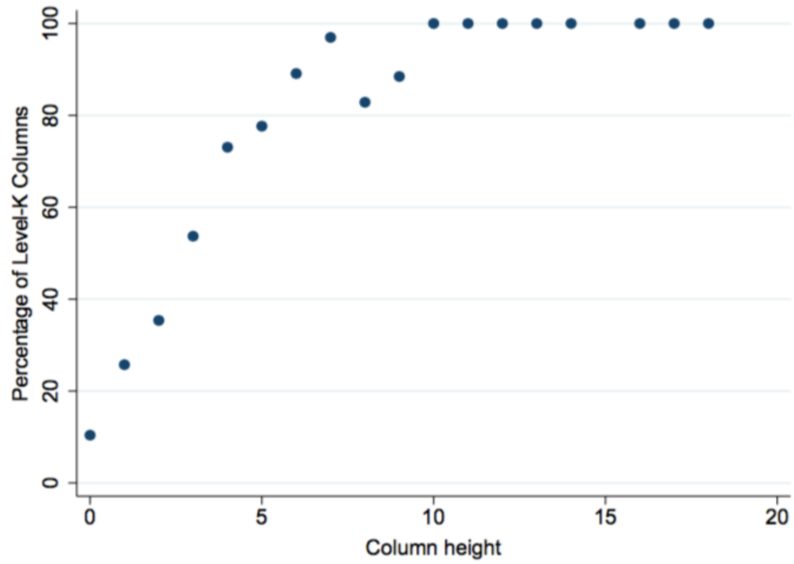
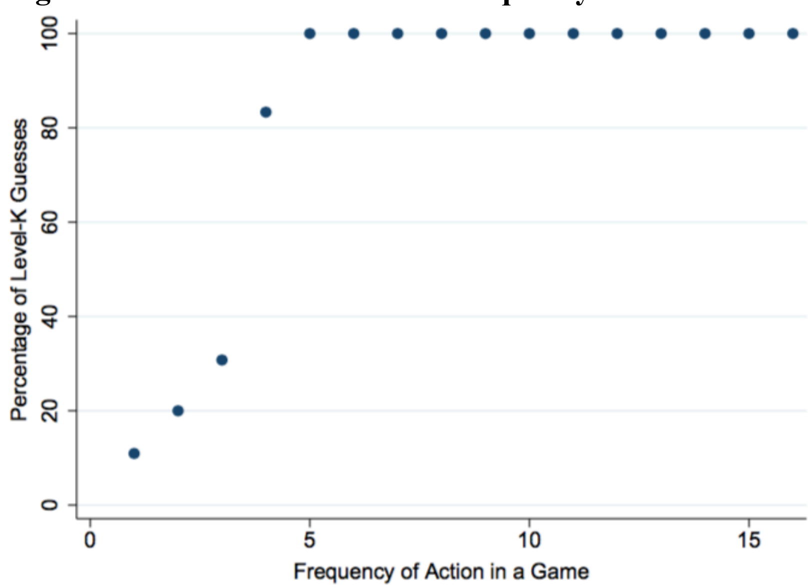


Figure 8: Percentage of Level *k* actions vs. action frequency



We obtain similar results when we consider behavior in the Actions treatment. Figure 8 organizes Actions behavior according to how many participants play the same strategy in a given NS game (x-axis). Empirically, we find that there are at most 16 subjects (out of 40) who select the same action in an NS game, *g*. The y-axis in Figure 8 shows the percentage of strategies that are constructed on Level k actions. The blue dot corresponding to 4 along the x-axis has a height of just over 80%, meaning that, when considering strategies that are played in a certain game by exactly 4 subjects, just over 80% of such strategies are Level k actions. Each dot with an x-value of 5 or greater has a y-value Level k frequency of 100% (Result 5). This suggests that Beliefs participants are not "missing" any types when they express their beliefs.

Result 5. If 5 or more Actions treatment participants select a given action in a Number Selection Game, g, then that action is a Level k strategy.

Another reason we expect U types do not believe in unknown deterministic types is that, on average, they distribute their boxes more in BA tasks compared to C subjects. For each Beliefs subject i, we compute s_i, the sum the squared heights of each column built.[23] The s_i values of C participants are greater than those for U subjects (Result 6).

Result 6. The mean sum of squared column heights for C and U participants are 875 and 444, respectively. This difference is highly significant using a two-sided t-test ($p < 0.0001$).

Despite the aforementioned results, it may be the case that U participants anticipate deterministic behavior, yet, may not express it because they are less confident in their beliefs and more risk averse in comparison to C subjects. We thus consider the responses by U and C participants in the Relative Performance (RP) questions and Bomb Risk (BR) decision. For a Beliefs subject i, we define her confidence level as the average fraction of subjects in i's session believed to have performed weakly worse than i in the BA tasks. (This is computed using the responses to the RP questions.) Thus, if i believes that, in each BA task, she has weakly more

[23] Using this measure has its advantages over "simpler" alternative statistics. For instance, if we were to simply count the number of columns subjects construct or the average height of each constructed column, we would not be able to differentiate between a participant who always places 19 boxes on one action and 1 on another from a subject who always places 10 boxes on each of two actions.

overlapping boxes than all others in her session, her confidence level is 1. If she believes that all others have strictly more overlapping boxes than does she in each BA task, her average confidence level is 0. We do not find any evidence that C participants are more confident in their constructed histograms in comparison to U subjects (Result 7).

Result 7. The average confidence levels of C and U participants are .700 and .667, respectively. This difference is insignificant using a two-sided t-test (p = 0.2112).

In both treatments, subjects' points are converted to monetary payoffs using binary lotteries (Roth and Malouf, 1979). This decision was to incentivize participants to maximize their expected number of points irrespective of their risk preferences. Binary lotteries do not always work in practice (see Selten et al., 1999), but they seem to have served their purpose in our study. In the BR decision, if we consider the average number of "treasure chests" opened by a Beliefs participant, we find that U subjects do not open significantly fewer chests than C participants. In other words, U subjects are not significantly more risk averse than C participants (Result 8).[24]

Result 8. The average numbers of boxes opened by C and U participants in the BR decisions are 49 and 46, respectively. This difference is insignificant using a two-sided t-test (p = 0.4206).

Given the findings up to this point, the next question to address is whether U subjects are creating flatter histograms than C participants because they (i) believe Actions participants are noisy or (ii) are noisy themselves. If (ii) is a more accurate explanation, it seems reasonable to

---

[24] More generally, we see that the binary lotteries used in the BA tasks seem to have effectively neutralized risk attitudes among all Beliefs participants. An OLS regression of the sum the squared heights of each column built on the number of treasure chests opened in the BR yields a positive slope coefficient with p-value of 0.544.

expect U participants to be less confident in their choices compared to C subjects (who we know are not noisy). We know, however, from Result 7, that U subjects are no less confident in their histograms than C participants. It thus seems that (i) is more plausible than (ii). In fact, simply looking at the histograms made by U subjects suggests more evidence supporting (i): it appears that U subjects arrange their boxes more towards the upper bounds of NS games in comparison to random allocations.

To investigate this more rigorously, we compute the following "front-load factor" ${f_i}^h$ for subject i's histogram *h*. Specifically, if *h* corresponds to a game with upper bound UB, then ${f_i}^h = \sum_{j=1}^{UB} h_j$, where h_j is the total number of boxes placed across actions *UB−j+1* through UB in h. Thus, if h has all 20 boxes on the upper bound, ${f_i}^h = 20 \times UB$. If *h* has 19 boxes on UB and 1 box on UB−1, then ${f_i}^h = 19 + 20 \times (UB - 1)$. The smallest ${f_i}^h$ can be is if *h* has all 20 boxes placed on its lower bound, leaving ${f_i}^h = 20$. For Beliefs participant *i*, we average her front-load factors over all 11 BA tasks to compute her "front-load statistic", $F_i = ({f_i}^1 + \cdots + {f_i}^{11})/11$. We find that U participants front-load their histograms more than would random subjects (Result 9). Furthermore, C subjects front-load their histograms more than U participants (Result 10).

Result 9. The mean front-load statistic for U participants is 254 which is significantly differ- ent than 221 ($p = 0.0126$), the expected front-load statistic reached via random box allocation.

Result 10. The average front-load statistics for C and U participants are 323 and 254, respectively. This difference is significant using a two-sided t-test ($p < 0.0001$).

2.3.3 Stepwise reasoning is more than "as if" theory of beliefs

In addition to obtaining Beliefs participants' choices, we record the order in which they arranged their boxes in the BA tasks, allowing us to examine how individuals transition between

Level k and k' actions for $k, k' \in \{0,1,2,3,4\}$. Because the instructions focus on explaining the payment method for overlapping boxes and the order of the arrangement of boxes is not incentivized, recording the order is minimally invasive. We consider all 25 (5 × 5) pairs of transitions, determining whether a box placed in one category predicts where the next box will be placed.

To do so, we construct several 5×5 transition matrices. The first is the empirical transition matrix, E, whose (i, j) entry corresponds to the number times two consecutive boxes are placed first on the Level $i-1$ strategy and then on the Level $j-1$ action. The second is an analogously defined random matrix, R, that assumes Beliefs participants arrange their boxes in random orders (but still produce the same final empirical histograms). The third is the normalized transition matrix, N, where N = E – R. [25]

Table 4 shows the matrix N. Below element $N(i,j)$ we report the p-values of a one- tailed binomial distribution test of the null hypothesis that $N(i,j) = 0$ given the final empirical distribution. Cells are colored if $N(i,j) > 0$ and has $p < 0.01$. First, we note that the diagonal cells (in blue) are all positive and significant, indicating that a box placed in a category is likely to be preceded by a box in that same category (Result 11).

[25] A subject is allowed to add and remove boxes without any restrictions. When we record the order in which boxes are placed, we only consider the 20 boxes that make up the histogram that is finally submitted.

Table 5: Level *k* action frequencies in Actions Treatment

	L0	*L1*	*L2*	*L3*	*L4*
L0	1010.9 (0.000)	171.4 (0.000)	-145.5 (0.000)	-106.6 (0.000)	-50.9 (0.000)
L1	-70.6 (0.000)	1032.8 (0.000)	97.5 (0.000)	-108.5 (0.000)	-54.2 (0.000)
L2	-109.3 (0.000)	-124.3 (0.000)	490.1 (0.000)	130.5 (0.000)	-30.9 (0.000)
L3	-75.8 (0.000)	-99.3 (0.000)	-21.5 (0.005)	170.7 (0.000)	86.5 (0.000)
L4	-43.2 (0.000)	-56.3 (0.000)	-27.2 (0.000)	6.7 (0.089)	64.4 (0.000)

Result 11. For all nine categories, if a box is placed in a category, C, the following box is likely to be placed in C (see the blue cells in Table 5).

If Step *k* thinking is more than an "as if" representation of behavior, we would expect individuals to express Level *k* beliefs before they express Level *k + 1* beliefs. The yellow cells in Table 4 show that transitions where a Level *k + 1* box is placed immediately after a Level *k* are significant for $k \in \{0,1,2,3,4\}$ (Result 12).

Result 12. For $k \in \{0,1,2,3\}$, if a box is placed on a Level *k* action, the following box is likely to be placed on the Level *k + 1* action (see the yellow cells in Table 5).

2.4 Related Literature

The Number Selection (NS) games are inspired by a variety of existing games. They closely resemble the Generalized Centipede (GC) games from Fragiadakis et al. (2017). In GC games, players select integers from (possibly different) guessing ranges. As with our NS games, guessing *x* guarantees a player earns at least *x*. In addition, there are bonuses that can be attained for undercutting as well as matching one's opponent. The GC games are less constrained than ours. We impose more restrictions on ours for parsimony, ease of explanation to experimental subjects and to facilitate our analysis.

The GC games were preceded by the 11-20 Money Request (MR) game studied in Arad and Rubinstein (2012). In the two-person MR game, each player simultaneously selects an integer between 11 and 20 (inclusive). Guessing x earns a player x, unless x is exactly 1 less than the opponent's guess, in which case the player earns $x + 20$. The Level 0 prediction is 20, the upper bound. Then, Level 1 is 19, Level 2 is 18 etc. The MR game has inspired others as well. For instance, Georganas et al. (2015) and Goeree et al. (2013) and Alaoui and Penta (2016) study games that are very similar to the 11-20, game, except for some changes in the payoff structures.

While we design our NS games, a more prominent feature of our design is our method for belief elicitation. A variety of scoring rules exist for eliciting beliefs; see Selten (1998) for an overview. The Quadratic Scoring Rule (QSR), for instance, is quite common.[26] As an example of QSR, consider a game with three pure strategies, a, b and c. When an individual reports A, B and C for the likelihoods that her opponent will play a, b and c, respectively, her payoffs are computed as $P - (A - 1\{s = a\})^2 - (B - 1\{s = b\})^2 - (C - 1\{s = c\})^2$, where $P > 0$ is a prize and $1\{s = x\}$ is an indicator function that equals 1 if and only if $s = x$ and equals 0 otherwise. It can be shown that a risk neutral agent is incentivized to truthfully express her beliefs to this mechanism.

Though QSR has been widely implemented,[27] data suggests it may have some

[26] The quadratic scoring rule was first studied by Brier (1950) and Good (1952).

[27] Costa-Gomes and Weizsacker (2008) use it to investigate whether a subject's actions in a normal form game is a best response to her stated beliefs of how others will play that game. Dufwenberg and Gneezy (2000) elicit beliefs in a Lost Wallet game to investigate how beliefs affect trust dilemmas. Dominitz and Hung (2009) explore how beliefs

shortcomings in practice. For example, Palfrey and Wang (2009) find that QSR elicits more extreme beliefs than an alternative (logarithmic) scoring rule that is also incentive compatible. Interestingly, Huck and Weizsacker (2002) find beliefs to be biased towards 50-50. Armantier and Treich (2013) highlight some additional drawbacks of incentive compatible scoring rules, namely, that stakes, incentives and hedging opportunities can substantially distort reported probabilities.

We believe there are two features of QSR that may pose some difficulty for real-world participants. First, it does not seem to convey its incentive properties transparently. We believe the design of our Box Arrangement tasks clearly do by making the incentives more visual via overlapping histograms. After the design of our experiment, we discovered that Carpenter et al. (2013) also use histograms to elicit beliefs, perhaps because they held similar concerns regarding a subject's understanding of the belief-elicitation mechanism.

Second, QSR asks a subject to think about how likely it is for a single individual to take various actions, akin to asking an individual for the probability that it will rain today. We believe an easier question would be: "how many days this week do you expect it will rain?" This is why we ask subjects to predict how many others choose different strategies as oppose to matching an individual with another and asking her for the likelihood that her opponent will choose different actions. Huck and Weizsacker (2002) also ask subjects to forecast quantities of other participants. In addition to the potential difficulty in formulating a probabilistic belief over a single occurrence (i.e., whether it will rain today), we think that asking an individual to predict the behavior of a single participant may trigger one to think extremely. In other words, if asked

are updated as new pieces of information are publicly announced. See Palfrey and Wang (2009) for a discussion of additional papers that have used QSR.

for the beliefs of whether an opponent will select pure strategy x or y, one may simply report the action that she believes her opponent is more likely to choose, which can explain the prior instances where QSR has recorded extreme responses.

2.5 Conclusion

This paper contributes to the behavioral game theory literature by testing whether the Cognitive Hierarchy (CH) model developed by Camerer et al. (2004) is more than simply an "as if" theory of behavior. CH describes an individual as believing her opponents engage in heterogeneous steps of strategic thinking. To test for such beliefs, we first have subjects in one treatment play a series of games. Their behavior is then predicted by a separate group of participants in a different treatment. As a participant builds a histogram to express her beliefs, she first anticipates a number of non-strategic individuals who play a naive strategy. Then, she believes that there are players who play the best-response to this strategy, followed by players who best respond to these strategic individuals and so on.

Our data show that CH cleanly passes the test of its beliefs assumptions, which may shed some light on a few existing puzzles in behavioral game theory. First, even in games whose most commonly selected actions are Level k strategies, a substantial portion of behavior is often unexplained by Level k; such behavior may, however, be rationalized by a Step k thinker best responding to some distribution of lower step thinkers. Second, the classification of an individual as a Level k player in a certain type of game has limited predictability as far as their behavior in other types of games.[28] A Step 2 thinker who believes $\alpha \in (0,1)$ and $1-\alpha$ of the population are Step *0* and Step *1* thinkers, respectively, may have a best response that coincides with the Level

[28] See Georganas et al. (2015). Similar results are found in Fragiadakis et al. (2017).

1 strategy in certain games but with the Level *2* strategy in others.[29] Lastly, our relatively "clean" results from our Beliefs treatment provide a proof-of-concept that our belief-elicitation method is not only theoretically appealing, but also practically successful. We hope future researchers interested in recording beliefs seriously consider our method; we expect this would help them to similarly obtain minimally noisy responses.

[29] Furthermore, she may adjust her beliefs of α depending on her perception of games' degrees of complexity.

3. MEASURING TRUST: A REINVESTIGATION

3.1 Introduction

Glaeser et al. (2000) (henceforth GLSS) investigate two ways of measuring trust. The first is an incentivized measure – the percentage of endowment sent in a variant of Berg et al. (1995) "investment game." The second is a survey-based measure, which consists of the widely used General Social Survey (GSS) trust question,[30] or a trust index.[31] An important finding in GLSS is the lack of correlation between the trusting behavior in the game and the survey measures of trust. Surprisingly, the authors instead find a positive relationship between the survey measures of trust and trustworthiness. They conclude that even though trusting behavior seems to have a stable individual-specific component,[32] it is at best weakly measured by typical attitudinal questions about trust. In this paper, we re-visit the investigation presented in GLSS by following their protocol. We conduct two experiments: a replication and a reinvestigation of GLSS. In both of the experiments we implement the original GLSS protocol. However, in the reinvestigation experiment, we introduce one major change: we employ the original Berg et al. (1995) investment game instead of the modified version used in GLSS. In contrast to GLSS and the replication experiment, in the reinvestigation, we observe a positive and significant correlation between trusting behavior in the lab and the responses to the survey questions about trust.

The game that GLSS use differs from the original investment game in two ways: the GLSS game endows only the first mover instead of both players, and uses a multiplier of two

[30] "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"
[31] The index combines the GSS trust question with two additional questions: "Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?" and "Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?"
[32] The authors show that the experimental measure of trust is correlated with self-reported trusting behavior and survey questions about trust towards strangers.

instead of three. The literature suggests that the asymmetry in the endowments may change motives for both players in the game (Ciriolo, 2007; Yan and Miao, 2007; Xiao and Bicchieri, 2010; Johnson and Mislin, 2011). The first mover may be motivated by altruism or inequality aversion, in addition to the trust motive. That is, the first-mover's decision to send money to an un-endowed second mover might reflect not only her trust in the second-mover, but also a desire to rectify the inequality in their initial endowments. The second mover, in turn, may return a smaller fraction of the amount received, also in an attempt to equalize payoffs. This change in the second mover's behavior may affect first mover's strategic beliefs, thus, distorting trusting behavior further.

In their meta-analysis, Johnson and Mislin (2011) studied the impact of various parameters of the investment game on trusting and trustworthy behavior. They reported that, when only the first mover is endowed, the amount sent by the first mover is higher. They argued that inequality aversion might create feelings of stress and guilt, when players are asymmetrically endowed. Thus, sending more to the second mover can help eliminate these feelings. Indeed, Johansson-Stenman et al. (2013) argued specifically that altruism or inequality aversion might explain the weak relationship between the survey-based trust measure and trusting behavior when the second mover is not endowed. Similarly, the second mover's apparent trustworthiness may also be impacted by the fairness motives as she claims her "fair" share of the total payment. [33]

Thus, the asymmetric endowments in the GLSS game may incentivize first movers to send more and second movers to send less than otherwise would be observed in the standard game, all else equal. However, if first movers anticipate the lower level of reciprocity, then the

[33] Yan and Miao (2007) show how the asymmetry in endowments can change trust and reciprocity behavior in trust game

trust component of their transfer may be reduced, and this could have an offsetting effect on the amount sent. Indeed, we find that while first movers send a slightly higher amount when endowments are asymmetric, the difference is not statistically significant. However, we do find that the percentage returned by second movers is significantly lower when endowments are asymmetric.

The difference in multipliers between the two versions of the game is a relatively minor consideration. The multiplier of three in the original game has a slight advantage in that it allows us to distinguish between different rules-of-thumb that might be used by second movers: paying back the entrusted amount, or splitting the total transfer received.[34] Our reinvestigation experiment reveals that the main motive behind trustworthy behavior is equalizing earnings rather than paying back the entrusted amount. If the first movers share the same motivation, their trusting behavior in the game with asymmetric endowments may be confounded with inequality averse preferences.

The weak relationship between the survey and the game measures of trust found in GLSS has been replicated in a study in Brazil, using the same modified investment game (Lazzarini et al. 2005). On the other hand, Johnson and Mislin (2012) found a positive cross-country correlation between the answers to the GSS trust question and average behavior in the investment game using a large data set of 80 investment game studies and country-level average survey question responses. Due to the nature of their data set, however, they were unable to analyze the relationship between the two measures of trust at an individual level. Additionally, they did not control for the potentially different versions of the investment games used in the investigated studies (asymmetry in endowments, for example). Using the trust game with

[34] This is true whenever the first mover sends all of her endowment and the second mover returns half of the doubled amount. This can be seen in Figure 1. See Ciriolo (2007) for a more in-depth analysis on this.

symmetric endowments, Capra et al. (2008) found that, at an individual level, after controlling for altruism, attitudinal questions are good predictors of trusting behavior. This suggests that the measures of *trusting behavio*r may capture both trust and the impact of social preferences on behavior. An accurate behavior-based measure of trust should, thus, minimize the impact of other social preferences on trusting behavior. The modified investment game used in GLSS may fail to achieve this objective as discussed below.

The main contribution of GLSS was to expose the lack of correlation between trusting behavior in the investment game and the survey measures of trust. This important finding has discouraged many researchers from using the survey questions regarding trust. Instead, economists have sometimes replaced the survey questions with incentivized games, which are costly and especially complicated to implement in the field studies (see Fehr et al. (2003) for an example; Wilson and Eckel (2011) address this issue in political science, where survey-based research dominates the study of trust). We find that the survey and the investment game measures of trust are correlated, when the modified version of Berg et al. (1995) investment game is used. Thus, our findings suggest that both the survey questions and the incentivized game may be used to measure trust.

3.2 Experimental Design

The experimental design and procedures of both of our experiments follow the GLSS protocol as closely as possible. While we replicate the GLSS protocol in the replication experiment, in the reinvestigation experiment we employ the original investment game with equal endowments (see Table 6 for details). Subjects were recruited using ORSEE (Greiner 2015) at Texas A&M University in April 2015 and October 2017 for the reinvestigation and the replication experiments respectively. Subjects completed an online survey one to seven days before the experiment. The survey contained demographic and attitudinal questions about trust and trustworthiness from the GSS as well as questions designed by GLSS. A total of 203 students completed the survey and 154 of these participated in the experiment in the Economics Research Laboratory at Texas A&M University. The subjects who completed the survey but did not participate in the experiment are not statistically different from those who did (see Table A1.1. in Appendix 1). This holds true in both the replication and the reinvestigation.

Table 6: GLSS, replication and reinvestigation design comparison

	GLSS	**Experiment 1: Replication**	**Experiment 2: Reinvestigation**
Survey Payment	Not Reported [a]	$5	$5
Show-up fee	Not Reported [a]	$5	$5
Recruitment	Classroom	ORSEE	ORSEE
Survey	Pen and Paper	Online	Online
First Mover Endowment	$15	$15	$10
Second Mover Endowment	$0	$0	$10
Multiplier	2	2	3
Subjects	Harvard Students	Texas A&M Students	Texas A&M Students
Number of Subjects	189	88	66

[a] We consulted with one of the authors, who was unable to recall the details about the survey payment and the show-up fee.

Following GLSS, subjects were paired in the order of arrival. They were seated in pairs and given 10 minutes to fill out a social connection survey, together with their paired partners. After completing the survey, they returned to their assigned seats to continue the experiment.

Each experiment includes a baseline investment game (no-promise condition) and a treatment that additionally allows for the second mover to make a non-binding promise to return as much as is sent to her or not (promise condition). The promise is made by checking one of the two boxes on a sheet of instructions before the game starts. This sheet is displayed to the first movers before they make any transfer decisions. 98 out of 196, 40 out of 88, and 34 out of 66 subjects are in the promise condition in the original, replication and reinvestigation studies respectively. Surveys, instructions and the script are presented in Appendix 3.[35]

3.3 Results

The analysis follows that presented in GLSS, which is reproduced using the original data provided by the authors.[36] Table A1.2 in Appendix 1 compares the original and our subject groups in detail. The table shows that there are significant differences between our sample and the GLSS sample. Our participants are more likely to be male and only children, less likely to be white, or to be freshmen. The subjects in our study show a slightly lower score on the GSS trust question, but not on the overall index.

Summary statistics of the results for both studies are presented in Table 7. In the original study, the percentage of endowment sent by the first movers (a measure of trust) is five percentage points higher than that in the replication experiment (Wilcoxon rank-sum test *p=0.03*). This is consistent with the idea that our subject pool is overall slightly less trusting. Although the percentage of endowment sent in the replication experiment is in turn higher than

---

[35] Like GLSS, we also conducted a second experimental measure of trust: the Envelope Drop task (not reported here). Students were asked to report their valuations for an envelope containing $10 that is addressed to the subject, and may be dropped at various public places in town. We did not find any significant effect of the reservation values calculated using the Envelope Drop data in any of the regressions, and so we chose not to present the data in this paper. The materials are available on request.

[36] All variables are constructed as in the original study (see Appendix 2 Table A2.1). We are able to reproduce the regression results using the data provided by the authors except for a few minor discrepancies.

the one in the reinvestigation experiment, where there are equal endowments, this difference is not statistically significant (Wilcoxon rank-sum test *p=0.31*).

The return ratio (a measure of trustworthiness), is calculated by dividing the amount returned by the doubled amount sent (tripled amount sent in the reinvestigation). In the original study the return ratio is five percentage points larger than that in the replication experiment, and the difference is statistically significant (Wilcoxon rank-sum test *p=0.01*). In the reinvestigation experiment, the return ratio is significantly larger than that in replication experiment (Wilcoxon rank-sum test *p=0.00*), or the original study.

Table 7: Summary results

	GLSS	**GLSS vs. Replication**	**Replication**	**Replication vs. Reinvestigation**	**Reinvestigation**
No Promise					
% Sent (First Mover)	83% (30%)	> * (0.03)	77% (27%)	= (0.31)	72% (32%)
Return Ratio	45% (26%)	> ** (0.01)	40% (24%)	< ** (0.00)	55% (20%)
N	189		88		66

Note: In GLSS and *Experiment 1*, only the first movers are endowed with $15 and the multiplier is 2. In *Experiment 2*, both movers are endowed with $10 and the multiplier is 3.

Standard deviations are in parentheses. *p<.05; **p<.01

In testing for treatment effects between the no-promise and the promise conditions, we find no significant treatment effects in the percentage of endowment sent or the return ratio for any of the three experiments (Wilcoxon rank-sum test *p=0.58, p=0.34, p=0.44, p=0.65, p=0.76 and p=0.25* respectively; details available on request). Thus, we pool the treatment and control data together for the remaining analysis.

These summary statistics suggest that using the standard investment game changes the behavior of the second mover, but not necessarily the first mover. Johnson and Mislin (2011) report that endowing both movers decreases the amount sent by the first mover, however, we find no significant difference in the percentage of the endowment sent. Nonetheless, the motives behind trusting behavior may change. For example, inequality averse individuals may be incentivized to send less, when both movers are equally endowed. However, they may also be incentivized to send more, if they believe equally-endowed second movers will send more money back. The two effects may offset each other.

The difference in the trustworthy behavior between the replication and the reinvestigation is consistent with Ciriolo (2007), who shows that a large proportion of the second movers employ a norm to equalize earnings. Thus, symmetric endowments mean that second movers will send a larger proportion of the multiplied amount in order to equalize earnings. This finding is also supported by Figure 11.

Figures 9, 10 and 11 illustrate the relationship between the percentage of the endowment that is sent by the first mover and the return ratio for the second mover for the original, replication and reinvestigation data sets, respectively. The sizes of the bubbles are proportional to the frequencies of the data points at each location. Subjects located on the solid line returned everything they received (13% of GLSS; 9% of replication; 3% of reinvestigation). The dashed line represents returning the exact amount sent by the first mover, as if they view the decision as "repaying the loan" by the first mover (57% of GLSS; 41% of replication; 15% of reinvestigation). Finally, the dotted line represents decisions that equal total earnings. Subjects located on this line returned an amount that equalizes final earnings for both players (55% of GLSS; 43% of replication; 55% of reinvestigation). [37] Since the second mover in GLSS and the replication was not endowed, we cannot disentangle the motives for the 44 and 18 pairs respectively where the first movers sent everything ($15) and the second movers returned 50% of the amount that was available to return ($15). The motive for these second movers could have been equalizing the earnings or returning the amount sent by the first mover. In our reinvestigation experiment, however, we observe that the dominant motive for the second movers was to equalize earnings, since observations are closely clustered around the line indicating equalized earnings.

Table 8 examines whether the answers to the GSS questions and other individual characteristics correlate with trust. The dependent variable is the amount sent by the first mover. The first two regressions use the original data, the second two use our replication data and the last two use our reinvestigation data. The six control variables (opposite gender pair dummy, promise condition dummy, gender dummy, white sender dummy, freshmen sender dummy and sender with no siblings dummy) do not correlate with the trusting behavior in the original or the replication data. In the reinvestigation data, the only significant control variable is the opposite gender pair dummy, which carries a significant positive sign.

---

[37] In GLSS, since the second mover was not endowed and multiplier was 2, the equal earnings line has a different equation ($y = 1.5(x - 5)$), than the one for our data ($y = 2x$), where y is the amount returned and x is the amount sent. Ciriolo (2007) offers an in-depth analysis on this.

Figure 9: The relationship between percentage of endowment sent and percent returned of the multiplied amount sent (GLSS)

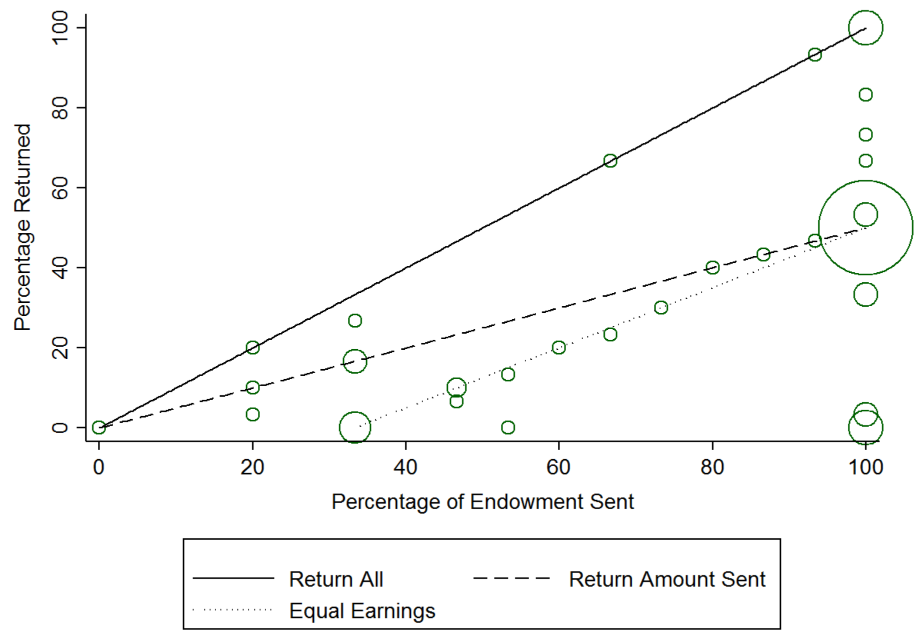


Figure 10: The relationship between percentage of endowment sent and percent returned of the multiplied amount sent (replication)

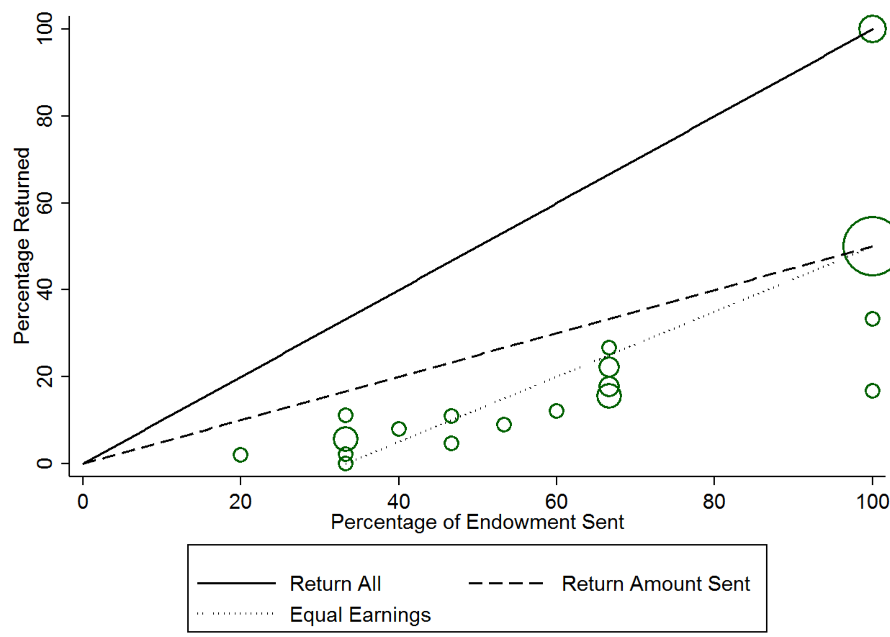
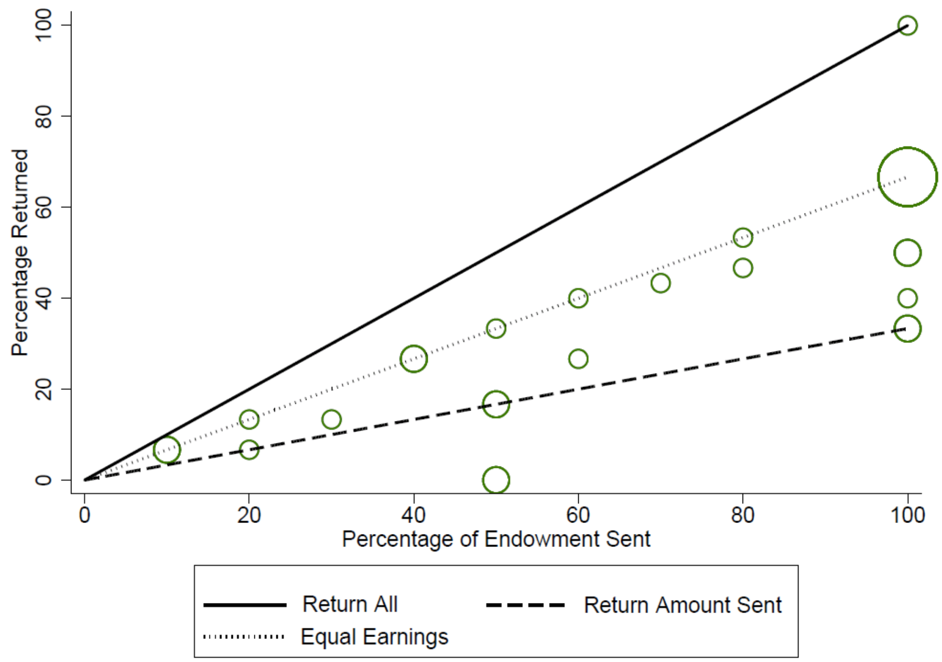


Figure 11: The relationship between percentage of endowment sent and percent returned of the multiplied amount sent (reinvestigation)



We replicate all of the original results regarding trusting behavior in the replication experiment and most of those results in the reinvestigation. In the reinvestigation, however, unlike the original or the replication, we find that both the single GSS trust question and the index of three related questions are positively and significantly correlated with the trusting behavior in the game. Subjects, who agree with the statement, "most people can be trusted," send on average \$2.70 more than the subjects who think "you cannot be too careful in dealing with people."

Interestingly, GLSS shows that answers to the GSS trust questions are correlated with trustworthy rather than trusting behavior in the lab. Unlike the original study, we do not find any significant correlation between trustworthiness and the survey measures of trust in the replication as well as the reinvestigation. Tables A1.3.1 and A1.3.2 in Appendix 1 present these results.

Table 8: Amount sent as a function of sender characteristics

	GLSS		Replication		Reinvestigation	
	(1)	(2)	(1)	(2)	(1)	(2)
Opposite gender	-0.670 (1.130)	-0.128 (1.112)	0.693 (1.321)	0.702 (1.323)	3.495*** (1.027)	3.171*** (1.050)
Promise treatment	0.043 (1.024)	-0.097 (1.015)	0.0189 (1.297)	0.057 (1.294)	-0.029 (1.039)	-0.452 (1.061)
Male	0.147 (1.197)	0.623 (1.174)	-1.715 (0.130)	-1.813 (1.388)	0.514 (0.969)	0.392 (1.008)
White	-0.330 (1.030)	-0.640 (1.025)	-1.840 (1.372)	-1.871 (1.372)	-0.720 (1.062)	-0.359 (1.082)
Freshman	-0.205 (1.136)	-0.434 (1.125)	1.015 (1.343)	1.149 (1.401)	-2.041 (1.559)	-2.140 (1.625)
Only child	-1.620 (1.531)	-1.724 (1.474)	2.516 (2.225)	2.461 (2.226)	0.282 (1.031)	-0.263 (1.069)
GSS trust	0.220 (1.022)		0.517 (1.352)		2.711** (1.104)	
Trust index		-0.094 (0.222)		0.480 (0.493)		0.450* (0.240)
Constant	12.920*** (1.76)	13.01*** (1.74)	12.271*** (1592)	12.429*** (1.544)	4.461*** (1.383)	5.905*** (1.288)
Adj.R^2	-0.059	-0.050			0.2814	0.2179
Observations	93	90	44	44	33	33

Note: Our constant coefficients for GLSS are slightly different from the published version of the paper.
*Standard errors are in parentheses. Significance: * p<0.10, ** p<0.05, ***p<0.01*

3.4 Conclusion

This paper follows the GLSS protocol to replicate the GLSS result: the lack of correlation between the survey and the laboratory measures of trust. However, we also conduct a reinvestigation of GLSS by replacing the modified version of the investment game used in GLSS with the original Berg et al. (1995) game.[38] To compare to the GLSS results and the replication, in the reinvestigation experiment, we observe a significantly higher percentage returned by the second movers. This finding is in line with the literature that shows how asymmetric endowments change the behavior in the investment game (Ciriolo, 2007; Yan and Miao, 2007; Xiao and Bicchieri, 2010; Johnson and Mislin, 2011). We find evidence that the answers to the GSS trust questions are significantly correlated with trusting behavior in the lab, when the standard game is used. Additionally, in contrast to GLSS, we do not find any evidence that the GSS trust questions are correlated with trustworthiness in either the replication or reinvestigation experiments.

The behavior observed in the trust game most likely captures trust along with other-regarding preferences. However, we argue that other regarding preferences might be more pronounced in the presence of asymmetric endowments. This could explain the different findings between this paper and GLSS. We find supporting evidence for the discussion by Johansson-Stenman et al. (2013), who noted that altruism or inequality aversion might, in fact, explain the weak relationship between survey-based trust and trusting behavior when the second mover is not endowed.

The important study by GLSS showed that survey and incentivized measures of trust were not correlated, implying that the two tasks measured different underlying constructs. Our findings suggest that trust is one construct, and that either the survey questions or the incentivized game may be used to measure trust. Thus, we conclude that survey measures of trust may indeed be a reasonable, low-cost alternative to using time consuming game measures of

[38] We do not attempt to disentangle the effects of changing the endowment and the multiplier in the game. However, based on the existing research, we believe that the endowment is the consequential change (see Ciriolo (2007), Johnson and Mislin (2011), and Johansson-Stenman et al. (2013)).

trust.

4. NORM MISPERCEPTIONS AND SOCIAL NETWORK STRUCTURE

4.1 Introduction

A large body of research in Economics and Sociology investigates the effects of social norms on individual behavior: pro-social actions (Bnabou and Tirole, 2006), cooperation and punishment (Fehr and Fischbacher, 2004), health-related behaviors (Perkins, 2005) and even tax compliance (Wenzel, 2005). Due to the strong relationship between social norms and behavior, social norms are sometimes referred to as the “Vehicle for Social Change” (Tankard and Paluck, 2016).

There exist multiple definitions of a social norm. Cristina Bicchieri (2010), for example, identifies two elements of social norms: the empirical and the normative. In particular, social norms can be defined empirically in terms of behavior in a group or normatively in terms of preferences for behavior in a group. In this paper, we will refer to a social norm as the average behavior in a population. Some prior studies we refer to below use a different definition of a social norm. The results presented in this paper extend beyond our definition of a norm, however, assuming a norm itself is observable and equivalent to behavior simplifies the discussion presented here.

Even though individuals seem to exhibit a tendency to conform to social norms, it is unclear, how information about the social norms is acquired and how beliefs regarding the social norms are formed. In fact, studies investigating health behaviors, criminal activity and peer effects suggest that individuals do not possess accurate beliefs regarding the social norms prevalent in the population. Social norm misperceptions have been documented in a wide variety of behaviors: alcohol consumption (Perkins et al., 2005), bullying (Perkins et al., 2011), sexual health behaviors (Young et al., 2013), gambling (Larimer et al., 2003), reports of attractiveness

(Bergstrom et al., 2004), sustainable transportation use (Kormos et al., 2015), mobile advertising use (Soroa-Koury et al., 2010), and delinquency (Yound et al.). For example, Neighbors et al. (2016) demonstrate that a group of college students from three university campuses report to consume an average of 10.59 alcoholic drinks per week while the same students estimate that their peers consume an average of 15.27 alcoholic drinks per week. The difference of almost 5 alcoholic drinks may be interpreted as a social norm misperception. Such norm misperceptions in behaviors that may put individuals' well being at risk is a cause of concern provided the large influence the social norms have on behavior.

In the context of socially complementary behaviors, where own engagement in an activity depends on the level of activity of others, the smallest norm misperceptions may incur considerable welfare losses to an individual and the society as a whole. The consequences of norm misperceptions may include abuse of alcohol, cigarettes and drugs, lack of contraception use, and over-engagement in unhealthy eating habits (Young et al., 2013; Perkins et al., 2005). It is not clear, whether social norm misperceptions cause suboptimal socially complementary behavior. However, empirical research suggests that informing individuals of the actual population norm improves their estimates of the norm and, importantly, reduces potentially harmful behavior in the future. For example, Neighbors et al. (2016) introduce an information intervention - they inform college students of the actual number of alcoholic drinks their peers report to consume in a week. They observe that three and six months after the intervention, students in the treatment group, relative to those in the control group, report to consume fewer alcoholic drinks and estimate others to consume fewer alcoholic drinks as well. Similar interventions are used to reduce bullying (Perkins et al., 2011), improve sexual health behaviors (Young et al., 2013), eating habits (Bergstrom et al., 2004) and promote sustainable transportation use (Kormos et al., 2015). The positive effects of these interventions suggest that improvements in social norm perceptions may lead to more optimal behavior.

Despite of the interest in norm misperceptions in the empirical literature, the underlying mechanism behind norm misperceptions has not been thoroughly investigated. Understanding the mechanism behind the misperceptions may help improve behavior. In his recent paper Jackson (2017) suggests a potential mechanism for how these belief biases arise in the context of socially complementary behaviors. Jackson suggests that social norm misperceptions may arise due to the structure of social networks. Simply put, individuals with higher preferences for a socially complementary activity engage into more social interactions to extract higher returns from this behavior. As a result, they are observable to more individuals in the population. Thus, on average, nodes in the network observe disproportionately more high preference nodes than low preference nodes. If individuals directly extrapolate their observation of network neighbors' behavior to form beliefs about the behavior of the population, their beliefs are likely to be upwards biased. As behavior in question is socially complementary, an upwards biased perceived social norm encourages individuals to engage in the levels of activity over and above the optimal level.

Whether or not social norm misperceptions arise as suggested by Jackson (2017) is an empirical question. Lerman et al. (2016) analyze the environments in online social networks such as those in Facebook and Twitter to demonstrate that most of the social media users are subjected to environments, where social network structure may result in a "majority illusion". That is, less active social media users are connected to disproportionately many active media users. Such an environment fits in with the description of networks in Jackson (2017) - it may result in social norm misperceptions. Lerman et al. (2016) though do not measure any social norm perceptions or social behaviors to test, whether social norm misperceptions do, in fact, take place.

The goal of the work presented in this paper is to test, whether social norm misperceptions can be generated by the mechanism outlined in Jackson (2017), and to uncover, whether the misperceptions, if any, translate to suboptimal levels of complementary behavior. We use a laboratory experiment to test the mechanism in a context-free environment. Controlled laboratory environment provides an objective measure of behavioral change as opposed to the self-reported survey measures of behavioral change used in the studies that investigate social norm misperceptions in the field. The ultimate goal of the work presented here is to create a platform for testing interventions that could diminish social norm misperceptions. The experiment reported below fails to create a strong baseline for future studies of network-generated norm misperceptions. We discuss the potential shortcomings of our framework and suggest improvements for future work.

The outline of the paper is as follows: Section 4.2 presents the framework of the mechanism proposed in Jackson (2017), Section 4.3 introduces the experiment, Section 4.4 discusses the results and Section 4.5 concludes.

4.2 Potential Mechanism Behind Social Norm Misperceptions, Jackson (2017)

This section outlines the model of socially complementary behavior, provided by Jackson (2017). Jackson incorporates the existence of positive and negative externalities into his model, which we abstract from for simplicity. A subject maximizes one's expected utility U_i, subject to x_i, own choice of the level of socially complementary behavior, and d_i, the number of social connections one constructs in the social network (degree):

In Equation 4, θ_i represents an individual preference for behavior x, a is a social complementarity parameter, $E(x_{j\neq i})$ is the expected behavior of others in the population (perceived social norm), $c_x x_i^2$ is the cost of the chosen level of behavior and $c_d d_i^2$ is the cost of the chosen number of social links.

$$\max_{x_i, d_i} U_i(x_i, d_i) = \theta_i x_i + a x_i d_i E(x_{j\neq i}) - c_x x_i^2 - c_d d_i^2 \tag{4}$$

Jackson proves that the optimal choice of social links d_i, is increasing in preference θ_i. Thus, high preference agents are predicted to form more social links in the network than low preference agents. It is then easy to see that, given the choice of d_i, the optimal choice of the level of behavior x is:

$$x_i = \frac{\theta_i}{2c_x} + \frac{a d_i}{2c_x} E_i(x_{j\neq i}) \tag{5}$$

The optimal choice of x_i is, thus, an increasing function of own preference for the behavior, θ_i, and the number of own social network links d_i. The resulting equilibrium predictions are:[39]

$$E(x_{j\neq i})^* = E(x)^* = \frac{E(\theta)}{2c_x - aE(d)} \tag{6}$$

$$x_i^* = \frac{\theta_i}{2c_x} + \frac{a d_i}{2c_x} E(x)^* = \frac{\theta_i}{2c_x} + \frac{a d_i}{2c_x}\left(\frac{E(\theta)}{2c_x - aE(d)}\right) \tag{7}$$

[39] Jackson (2017) assumes an infinitely large population such that the correctly estimated value of $E_i(x_{j\neq i})$ is close to identical for all agents with finite consumption choices.

As a result, high preference agents are not only predicted to have more links in the social network but also select higher levels of behavior x. Consequentially, individuals in the social network observe disproportionately many high preference, high connectivity agents among their neighbors. If they anchor their beliefs about the behavior of the population at the observation of their social network neighbors' behavior, they form upwards biased beliefs $\widetilde{E_\iota}(x)^* > E(x)^*$. Given Equation 7, such beliefs lead to over-engaging in the socially complementary activity x.

4.3 The Experiment

4.3.1 Theoretical Model Versus Experimental Setting

The model of the socially complementary behavior described in Section 4.2 is a two-step process. Firstly, individuals form a social network. Given the fixed social network structure, agents make consumption decisions observable to their social network neighbors. To provide an initial test of whether social norm misperceptions can be generated in this setting, we fix the first step of the mechanism by exogenously imposing "social networks" upon a group of participants in the laboratory as outlined below.

In our networks, a link between two participants i and j means that agent i observes agent j's consumption choice x_j but the opposite does not have to be true. Thus, our networks are directional.[40,41,42] We will refer to the link, which indicates the information out-flow from j as an outward link, and to the link, which indicates the information in-flow to i, as an inward link. It is

---

[40] Directionless networks involve links between i and j, where i observes agent j's behavior and j observes agent i's behavior.

[41] Constructing directional networks leaves us freedom to easily extend the platform developed in this paper to allow for endogenous network formation. Constructing endogenous directionless networks would require making decisions about agreements that would constitute a link, which would complicate the experiment and the theoretical predictions.

[42] Twitter is a directional network while Facebook is generally a directionless network (ignoring features such the possibility to block the news feed from selected accounts).

important to note that Jackson (2017) does not explicitly discuss the directionality of the networks he describes. While the distribution of inward links may play an important role in information aggregation, in the context of his model, outward links are of key importance as they generate the bias in the emitted information. We will denote the number of outward links a participant i has as d_i, consistently with the notation in the model above. The same argument of the mechanism generating the social norm misperceptions applies in the directional and directionless networks; thus, we would expect to see the upwards belief and consumption bias in the setting described in Section 4.2 regardless of network directionality.[43]

4.3.2 Overview and Hypotheses

Our experiment involves a one-by-two design: a control and a treatment condition, which could be easily extended in the future (see Table 9), if successful.[44] We construct a control condition with unbiased networks, where all participants have the same number of outward links, in expectation, irrespective of their preferences. We also construct a treatment condition with biased networks, where individuals with higher preferences for the socially complementary behavior have more outward links than those with lower preferences, in expectation.[45, 46]

If individuals form social norm perceptions as suggested by Jackson (2017) such that they anchor their beliefs at the observed behavior of their social network neighbors ($E_i(x) =$

[43] Our experiment involves a finite sample of participants, however the directional predictions in the experiment are consistent with those presented in Section 4.2. Importantly, agents with higher preferences are incentivized to consume more than those with low preferences for any beliefs about the social norm $E(x_{j\neq i})$.

[44] Future extensions may involve investigating methods to reduce social norm misperceptions, studying network properties that facilitate or complicate information aggregation in networks, extending the test of the mechanism to allow for endogenous network formation, and others.

[45] Directionless networks involve links between i and j, where i observes agent j's behavior and j observes agent i's behavior.

[46] Constructing directional networks leaves us freedom to easily extend the platform developed in this paper to allow for endogenous network formation. Constructing endogenous directionless networks would require making decisions about agreements that would constitute a link, which would complicate the experiment and the theoretical predictions.

$\widetilde{E_i}(x)$), then the social norm misperception $\bar{E}(x) > E(x)^*$ should arise in the treatment condition, on average. Consequentially, in the treatment condition, we expect to observe a positive consumption bias $\bar{x} > x^*$ (see equation 7), on average. However, no such biases should occur in the baseline (subject to the concerns discussed in Section 4.3.8). Our hypotheses are summarized in Table 9 and outlined below.

Table 9: Overview of the experimental design

	Control (Unbiased Networks)	**Treatment (Biased Networks)**
Manipulation	$E\big(var(d)\big) = 0$ $E\big(corr(d,\theta)\big) = 0$	$E\big(var(d)\big) > 0$ $E\big(corr(d,\theta)\big) > 0$
Hypotheses	$\bar{E}(x) = E(x)^*$	$\bar{E}(x) > E(x)^*$
(assuming beliefs $\boldsymbol{E_i(x) = \widetilde{E_i}(x)}$)	$\bar{x} = x^*$	$\bar{x} > x^*$

The expected zero variance of degree in the control results in no Friendship Paradox.[47] *Thus, expected degree in the group is equal to the expected degree of neighbors $E(d) = \tilde{E}(d)$. This also implies zero degree-preference correlation. Even if participants anchor their beliefs about the group at the observed social network neighbors, they should, on average, form correct beliefs about the distribution of degrees and preferences in the group. The positive degree-preference correlation in the treatment condition should similarly generate upwards biased beliefs about $E(d)$ and $E(\theta)$, which would result in social norm misperception $\bar{E}(x) > E(x)^*$ and overconsumption $\bar{x} > x^*$.*

Hypothesis 1: Participants exhibit zero bias in beliefs and consumption in the control condition.

Hypothesis 2: Participants exhibit a positive bias in beliefs and consumption in the treatment condition.

---

[47] A property of networks with positive degree variance $var(d) > 0$, where the nodes' network neighbors possess more links than the nodes in the network themselves $E(d) < \tilde{E}(d)$.

Hypothesis 3: Participants exhibit a higher bias in beliefs and consumption in the treatment than the control condition.

4.3.3 Subject Types

Each subject in our experiment is assigned one of the two types: High or Low. High type players are induced with high preferences for an abstract socially complementary behavior x ($\theta_H = 60$) while Low type players are induced with low preferences for x ($\theta_L = 20$). High type players are also assigned degree $d_H = 2$ in the control and $d_H = 3$ in the treatment condition. Low type players are similarly assigned degree $d_H = 2$ in the control and $d_H = 1$ in the treatment condition. Changing the degree parameters create our main experimental manipulation as outlined in Table 9.

For $a = 3$, $c_x = 5$ and $c_d = 0$, subjects then earn points according to the type and condition specific points' functions that are equivalent to the expected utility maximization problem in equation 4:

$$\max_{x_i} Points_{H,c,i}(x_i) = 60x_i + 6x_i E_H(x_{j\neq i}) - 5{x_i}^2 \qquad (8)$$

$$\max_{x_i} Points_{L,c,i}(x_i) = 20x_i + 6x_i E_H(x_{j\neq i}) - 5{x_i}^2 \qquad (9)$$

$$\max_{x_i} Points_{H,t,i}(x_i) = 60x_i + 9x_i E_H(x_{j\neq i}) - 5{x_i}^2 \qquad (10)$$

$$\max_{x_i} Points_{L,t,i}(x_i) = 20x_i + 3x_i E_H(x_{j\neq i}) - 5{x_i}^2 \qquad (11)$$

where *H* denotes a High type player, *L* - Low type player, *c* - control condition, and *t* - treatment condition.

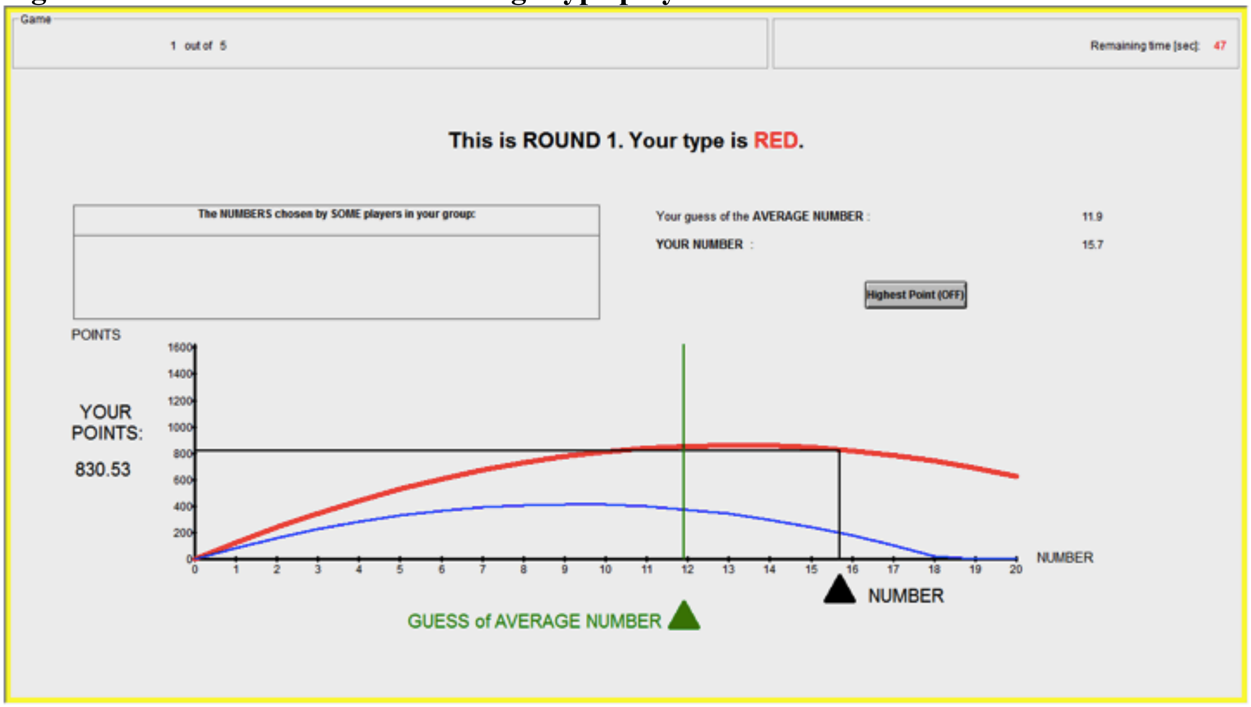
4.3.4 Decision Interface

In the experiment, subjects choose consumption levels $x_i \in [0,20]$ in 0.01 increments and estimate average consumption of others $E(x_{j\neq i}) -$ in 0.1 increments. Figure 12 illustrates the decision screen for a High type player. The decision problem is simplified by translating the mathematical form of the payoff functions to a visual interface. The interface is dynamic: the payoff functions change form as a subject submits different beliefs (points' curves shift to the right and rise as expected social norm increases). The payoff functions, thus, illustrate payoffs conditional on correct beliefs. Subjects are trained to use and understand the visual interface before they start the experiment.

4.3.5 States of The World

20 individuals participate in each session of our experiment. Within a session, subjects are randomly assigned to two groups of 10 participants. The groups remain fixed for the duration of the experiment. Subjects only interact with the members of their own group. One of the two 10-participant groups, selected at random, includes 6 High type subjects and 4 Low type subjects (High group), while the other includes 4 High type subjects and 6 Low type subjects (Low group) (see Table 10). Types of players are randomly assigned within each group. Each subject is, thus, equally likely to be a High or a Low type. Theoretically, the goal of each player is to identify the group one is assigned to - the randomly chosen state of the world. One then should be able to arrive at full-information Nash Equilibrium beliefs $E_i(x_{j\neq i})^*$and consumption $x_i{}^*$.

Figure 12: Decision interface for a High type player



The top left corner of the screen displays the game number. The top right corner of the screen displays the remaining seconds in the round. The text below the header of the screen displays the round number and subject type: High (Red) or Low (Blue). A subject selects consumption x_i *(Number) from 0 to 20 (on the x-axis) by moving the triangular NUMBER slider on the screen. The chosen NUMBER is displayed on the right hand side of the screen. A subject expresses one's belief* $E_i(x_{j\neq i})$ *by making a guess of the average Number (also, on the x-axis), by moving the triangular GUESS of AVERAGE NUMBER slider on the screen. Subject's GUESS of the AVERAGE NUMBER is displayed on the right hand side of the screen. The label on the y-axis displays the points that a subject earns for one's selected consumption* x_i *if one's belief* $E_i(x_{j\neq i})$ *is correct. High type players earn points illustrated by the red graph. Low type players earn points illustrated by the blue graph. A subject observes consumption levels chosen by one's social network neighbors in the table on the left side of the screen. Notice that, conditional on correct beliefs, the consumption choice problem is reduced down to selecting the highest point on the graph. This choice can be automated by clicking on the button "Highest Point", which forces consumption choice to be a best response to submitted beliefs about the social norm. If a subject forces best response, one's task is merely correctly estimating the average consumption of others by observing signals in the table on the left hand side.*

4.3.6 Random Network Formation

Upon arrival to the laboratory, every subject is assigned an identification number. Each player in the control condition possesses $d_i = 2$ virtual ID cards, while each High type player in the treatment condition possesses $d_H = 3$ virtual ID cards and each Low type player in the treatment condition possesses $d_L = 1$ virtual ID card. All subjects in the same group put their ID cards into a virtual hat. The computer then randomly draws 4 cards for each subject from one's group hat with replacement. A card is replaced with another draw if the drawn card has the subject's identification number for whom the card is drawn or if this subject already possesses a card with the drawn number. The 4 identification numbers drawn for player i represent i's 4 social network neighbors. Player i then observes his social network neighbors' consumption choices in the decision screen as illustrated in Figure 12. The network formation process is public knowledge to all participants.

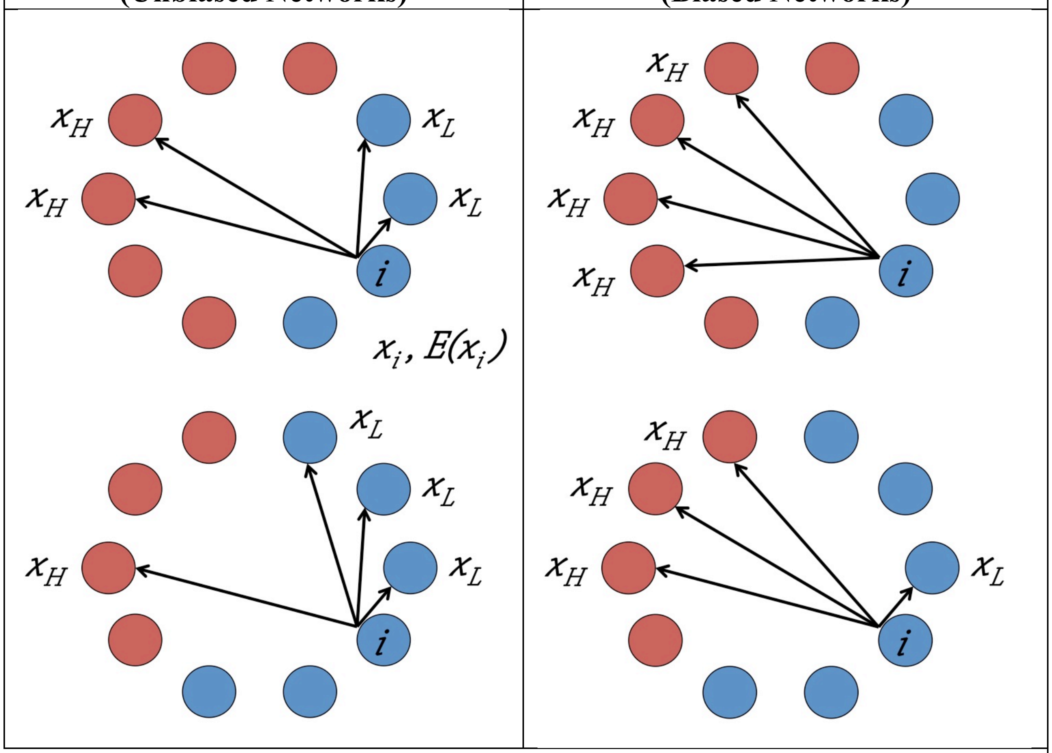
4.3.7 Experiment Timeline

Subjects in the experiment play five Games. In each Game, the state of the world for a group is randomly re-assigned: a group may be High or Low, subjects are randomly re-assigned a type: High or Low, and a new random network is implemented. Each Game consists of four Rounds. Each Round in Game 1 lasts one minute while each Round in Games 2-5 lasts thirty seconds. After each Game, subjects receive feedback about their points collected for that Game (see Table 11).

In each game, a subject makes two decisions at the same time. One selects the level of consumption and the expected norm of consumption. Both decisions are recorded for payment in one of the five randomly chosen games, at a randomly chosen point in time[48].

Table 10: States of the world in the control and treatment conditions

	Control (Unbiased Networks)	Treatment (Biased Networks)
High State		
Low State		



An individual observes disproportionately many high type (Red) players in the Treatment condition regardless of the state of the world. Thus, the average of the observed neighbors' consumption is a biased estimate of the average consumption in the group in the treatment, but not control condition.

4.3.8 Concerns of Information Aggregation

The emergence of norm misperceptions relies on the limited information aggregation in social networks. Theoretical literature provides some guidance as to when information aggregation in social networks takes place and when misinformation may spread. For example,

---

[48] The payment timing within a game follows an alpha beta distribution, consistent with Caplin (2011). Caplin demonstrates that such payment rule incentivizes a subject to make an optimal choice at each point in time.

Golub and Jackson (2010) demonstrate that if strongly connected networks grow large enough such that the influence of the most influential agent vanishes, a naïve learning rule[49] provides convergence to truth. Similarly, Jadbabaie et al. (2012) show that continuous arrival of information, where signals are independent over time but may be correlated within a time period among agents, ensures that a naïve learning rule leads to learning the true state of the world.

Table 11: Timeline of the experiment

Game	**Period**				
	1	**2**	**3**	**4**	
1	60s	60s	60s	60s	**Feedback**
2	30s	30s	30s	30s	
3	30s	30s	30s	30s	
4	30s	30s	30s	30s	
5	30s	30s	30s	30s	

The two slider bars in the decision interface in Figure 12 are placed at two randomly chosen positions on the screen at the beginning of each Game.

Our framework deviates from the two settings above as our networks are small and the signals about the state of the world are not only correlated across subjects but also across time. To provide the best chance for our subjects to form correct beliefs about the consumption in the group, we impose strongly connected networks[50,51].

[49] Naïve learning rule refers to averaging the signals about the state of the world received from one's social network neighbors.
[50] Strongly connected networks in which it is possible to reach any node starting from any other node by traversing edges in the outwards direction.
[51] Our networks are strongly connected with probability 0.99. This number represents the proportion of strongly connected networks among 100 randomly generated networks for parameter combinations $N = 10$ and $d_{in} = 4$.

One could alternatively rely on full information Nash Equilibrium predictions in Table 12, if one is willing to ignore the dynamic nature of information aggregation in the present setting.

4.3.9 Full Information Nash Equilibrium Predictions

The payoff functions in Section 4.3.3 give rise to the full information Nash Equilibrium predictions in Table 12.

Table 12: Full information Nash Equilibrium predictions

			Control (Unbiased Networks)	**Treatment (Biased Networks)**
High State	High Type	Beliefs	10.83	11.95
		Consumption	12.50	16.75
	Low Type	Beliefs	11.25	13.15
		Consumption	8.75	5.94
Low State	High Type	Beliefs	8.75	7.02
		Consumption	11.25	12.32
	Low Type	Beliefs	9.17	7.90
		Consumption	7.50	4.37

As the precise predictions vary by player type, state of the world and the treatment condition, our general hypotheses are summarized in Section 4.3.2. The bias in consumption and beliefs will be calculated by subtracting the relevant full information Nash Equilibrium predictions from the observed average consumption and beliefs respectively.

4.3.10 Procedures

Experimental subjects are undergraduates and graduate students, recruited using ORSEE (Greiner et al., 2015). Participants interact via a network of computers linked by z-Tree (Fischbacher, 2007) at the Economics Research Laboratory in TAMU's Department of

Economics. Two 20-subject sessions were run in the control condition, and three 20-subject sessions – in the treatment condition. Each session lasted approximately one hour. Average earnings are $26 including a $10.00 show-up payment.

4.4 Results

We firstly present the results inconsistent with our Hypotheses outlined in section 4.3.2. Table 13 summarizes the average bias across time and all participants, in their reported beliefs and consumption for both of the experimental conditions. We reject Hypothesis 1 as the bias in beliefs and consumption is positive in the control condition. Yet, we find support for Hypothesis 2 as the bias in the beliefs and consumption is positive in the treatment condition. However, subjects in the control condition exhibit a larger bias in consumption than the subjects in the treatment condition, which provides evidence against Hypothesis 3.

Table 14 presents analogical results for conditional belief and consumption bias where deviations are calculated from the actual average consumption in the group and the best response to these correct beliefs rather than from the full information Nash Equilibrium predictions outlined in Section 4.3.9. The treatment effect on conditional overconsumption vanishes, but conditional belief bias is higher in the treatment than the control group, which is in line with Hypothesis 3. The four tables in Appendix 4 also demonstrate that subjects do not seem to exhibit any convergence to optimal beliefs and behavior over time.

Our hypotheses, however, rely on several key assumptions that may fail in our experiment. The failure to comply with these assumptions may be the source of the bias in the control condition. We further investigate the assumptions of belief-action consistency and social learning dynamics in Sections 4.4.1 and 4.4.2. We additionally consider hedging, strategic player behavior and other obstacles to information aggregation in sections 4.4.3-4.4.5.

Table 13: Summary of results

	Control (Unbiased Networks)	Control vs. Treatment	Treatment (Biased Networks)
Belief bias $\bar{E}(x) - E(x)^*$	2.60 (2.43)**	<	2.72 (3.15)**
Consumption bias $\bar{x} - x^*$	2.51 (2.48)**	>*,**	2.18 (2.68)**

Results in this table are averages across a balanced panel presented in increments of one second.

Standard deviations are in parentheses. *p<.05; **p<.01.
Stars above the average bias statistics indicate significance of a one-sample t-test for a null hypothesis of zero bias.
Stars above the comparison signs indicate significance of a two-sample t-test for a null hypothesis of equal bias, and significance of a two-sample Wilcoxon rank-sum test.

Table 14: Summary of results for conditional bias

	Control (Unbiased Networks)	Control vs. Treatment	Treatment (Biased Networks)
Belief bias $\bar{E}(x) - E(x)^{*,actual}$	2.90 (3.11)**	<**,**	3.81 (3.65)**
Consumption bias $\bar{x} - x^{*,actual}$	2.23 (2.46)**	>	2.11 (2.56)**

Results in this table are averages across a balanced panel presented in increments of one second.

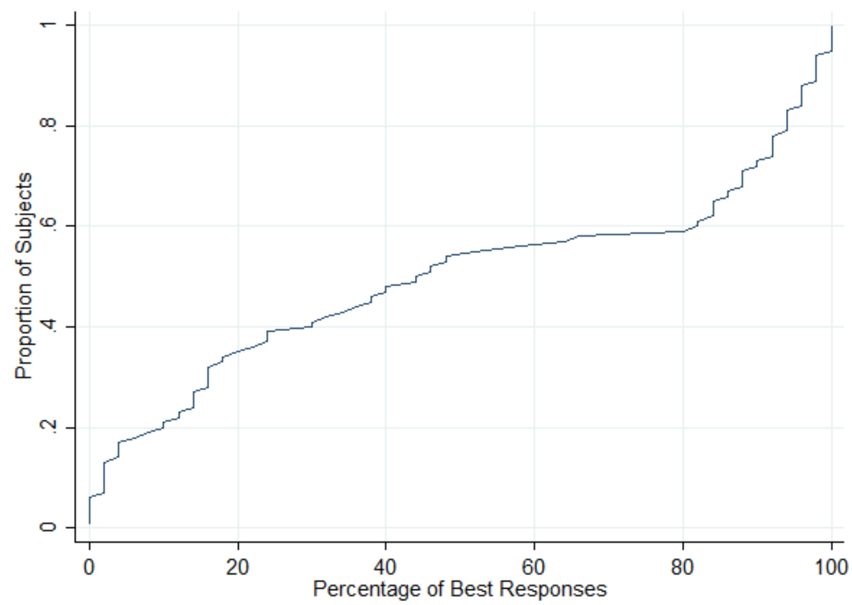
Standard deviations are in parentheses. *p<.05; **p<.01.
Stars above the average bias statistics indicate significance of a one-sample t-test for a null hypothesis of zero bias.
Stars above the comparison signs indicate significance of a two-sample t-test for a null hypothesis of equal bias, and significance of a two-sample Wilcoxon rank-sum test.

4.4.1 Belief and Action Consistency

One of the underlying assumptions in Jackson (2017) is individual belief-action consistency. Literature in behavioral game theory suggests that individuals may, in fact, choose actions that are not best responses to their reported beliefs (Costa-Gomes and Weizsacker, 2008). In our experiment, we observe, whether individuals best respond to their beliefs either manually by moving the slider arrows on the screen or by clicking the "Highest Point" button (see Figure 12).

We consider that a subject best responds to one's beliefs if one chooses consumption such that the difference between the chosen consumption and the optimal consumption given one's belief is at least as small as 0.5 units. Figure 13 depicts the cumulative distribution of the percentage of the best responses by subject in your experiment. It is apparent that approximately half of our subjects best-respond to their beliefs less than 50% of the time, while approximately 40% best-respond to their beliefs more often than 80% of the time. Consequentially, 62% of all our data is generated through a process of belief-action consistency.

Figure 13: Cumulative distribution of percentage of best responses by subject



Once we analyze only the data generated through belief-action consistency, we acquire the results presented in Table15. The directional effects in both beliefs and consumption bias are now consistent with Hypothesis 3. However, only the bias in beliefs is significantly larger in the treatment than the control group.

Consumption bias, that is not a best response to reported beliefs, is higher than that, which is a best responses to reported beliefs (1.36 < ** 2.44 and 1.84 < ** 2.16 in the control and treatment conditions respectively). Thus, the overall bias may be exacerbated due to the existence of players, who do not best respond to their beliefs. If the behavior of belief-action inconsistent players plays an important role in social norm misperception formation, it would be interesting to further study this seemingly irrational behavior. However, our experiment is not tailored for this specific purpose.

Table 15: Summary of results, conditional on best response to beliefs

	Control (Unbiased Networks)	**Control vs. Treatment**	**Treatment (Biased Networks)**
Belief bias $\bar{E}(x) - E(x)^*$	2.35 (2.11)**	<**,**	3.79 (3.11)**
Consumption bias $\bar{x} - x^*$	1.36 (1.30)**	<	1.84 (1.84)**

Results in this table are averages across a balanced panel presented in increments of one second.

Standard deviations are in parentheses. *p<.05; **p<.01.
Stars above the average bias statistics indicate significance of a one-sample t-test for a null hypothesis of zero bias.
Stars above the comparison signs indicate significance of a two-sample t-test for a null hypothesis of equal bias, and significance of a two-sample Wilcoxon rank-sum test.

4.4.2 Social Learning Dynamics

Jackson (2017) results also rely on an assumption that individuals form their beliefs about the behavior of the population by anchoring beliefs on the observed behavior of their social network neighbors. We investigate the prevalence of such belief formation in Table 16.

Results of the regression 1 in Table 16 indicate that for every one-unit increase in average network neighbors' consumption, an individual increases one's expected average consumption in the group by 0.25 units. This effect is significant across specifications 1-4, and 6. It persists, when we control for the control condition, high state of the world and high type indicators in specification 2. Regression 3 includes the proportion of best responses for a subject in the game, which has a positive effect on the expected consumption level in the group. Thus, subjects, who more likely to best respond to their beliefs, form higher expectations. From regression 4 it also appears that subjects do not merely respond to the average consumption of their network neighbors, they tend to respond particularly to the highest consumer among their network neighbors. Such an asymmetric following rule may, in general, produce the positive bias on its own. Regression 5 is run only for subjects, who exhibit less than 10% belief-action consistency, while regression 6 is run for subjects, who exhibit more than 90% belief-action consistency. It seems that only the more consistent individuals react to both the observed average consumption of neighbors and the overall average consumption in the group exhibiting very minor learning over time.[52]

The overall limited correlation between the average consumption of network neighbors and beliefs about the average consumption in the group may contribute to the exacerbated bias in both the treatment and control groups. Similarly, the lack of learning over time and the inability to correctly estimate the average group consumption may explain the positive bias in the control condition.

4.4.3 Hedging

Individuals may exhibit a positive bias in consumption in the low state of the world, if they hedge between the two states of the world. That, however, implies that a negative bias should be observed in the high state of the world. Table 17 breaks the summary results down by the state of the world.

---

[52] Note also that the constant in specification 5 in Table 16 is 10 units, which is exactly the middle of the strategy space. It may be that belief-action inconsistent individuals report random-like beliefs.

Table 16: Expected average consumption as a function of individual characteristics and experimental controls

Dependent variable: $E(x)$						
	(1)	(2)	(3)	(4)	(5)	(6)
Average consumption of network neighbors $\tilde{E}(x)$	0.25*** (0.04)	0.26*** (0.04)	0.26*** (0.04)	0.15** (0.06)	0.29 (0.26)	0.17** (0.08)
Average consumption of group members	-0.02 (0.05)	-0.02 (0.05)	-0.00 (0.05)	-0.03 (0.05)	-0.07 (0.19)	-0.23** (0.10)
Time (in seconds)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.001** (0.001)
Control condition		0.21 (0.23)	0.26 (0.24)	0.52* (0.28)	-1.84 (0.93)	-0.54 (0.38)
High state of the world		-0.14 (0.22)	-0.17 (0.22)	-0.08 (0.22)	-0.82 (0.83)	-0.33 (0.48)
High type		0.07 (0.22)	0.06 (0.22)	0.05 (0.22)	-0.24 (1.07)	-0.79* (0.44)
Proportion of best response per subject-Game			0.51* (0.30)	0.50* (0.29)		
Minimum consumption of network neighbors				0.03 (0.03)		
Maximum consumption of network neighbors				0.15*** (0.05)		
Constant	9.98*** (0.83)	9.63*** (0.89)	9.13*** (0.93)	7.88*** (1.03)	10.08** (4.68)	15.50*** (4.68)
R^2	0.04	0.04	0.04	0.05	0.20	0.08
Observations	101,450	101,450	101,450	101,450	3,923	19,365
*Standard errors are clustered at the subject-Game level. Significance: * p<0.10, ** p<0.05, ***p<0.01*						

Table 17: Average consumption bias by the state of the world

Belief bias $\bar{E}(x) - E(x)^*$	**Low state of the world**	**Low vs. high state of the world**	**High state of the world**
Control	3.53 (2.39)**	>**,**	1.49 (2.14)**
Treatment	3.40 (2.48)**	>**,**	0.96 (2.29)**

Results in this table are averages across a balanced panel presented in increments of one second.

Standard deviations are in parentheses. *p<.05; **p<.01.
Stars above the average bias statistics indicate significance of a one-sample t-test for a null hypothesis of zero bias. Stars above the comparison signs indicate significance of a two-sample t-test for a null hypothesis of equal bias, and significance of a two-sample Wilcoxon rank-sum test.

Table 17 shows that while the consumption bias is definitely larger in the low state of the world, it is still significantly positive in the high state of the world. Thus, it is unlikely that hedging could explain the overall observed positive bias in consumption.

4.4.4 Strategic Signaling

At the end of our experiment, participants completed a survey, where they were asked to report their reasoning behind the choices they had made in the study. 76% of the subjects recognized that they might have manipulated their group members' decisions through their own choices of consumption. 8 subjects explicitly reported to have chosen higher levels of consumption in order to push the average consumption level in the group up. Such behavior may have contributed to the average positive bias in the experiment. It also has an interesting interpretation in the context of undesirable socially complementary behaviors: for example, teenagers may use their social influence to peer pressure their friends into smoking or alcohol consumption by manipulating their perception of the social norm. Such strategic behavior may be very interesting to investigate, as it may prevent information aggregation in the social networks and may undermine the effectiveness of information interventions designed to change individual perceptions of the social norms. However, our experiment is not designed to identify the strategic motives in this setting.

4.4.5 Information Aggregation and Network Structure

As mentioned in Section 4.3.8, the influence of the most influential agent in the social network may be consequential to information aggregation. We measure the influence of an agent i in the social network in two ways: the number of actual outward links an individual possesses and betweenness[53] centrality. For each of the twenty unique networks implemented in the experiment we select the maximum number of outward links and the maximum betweenness centrality score as measures of the most influential agent. Figures 14 and 15 illustrate the relationship between the average bias and the maximum number of outward links in the fifty implemented networks. Figures 16 and 17 illustrate the relationship between the average bias and the maximum betweenness centrality score in the fifty implemented networks.

Figures below provide suggestive evidence that different measures of maximum influence may have a different effect on the observed belief and consumption bias. While maximum outwards degree is barely positively correlated with the bias in both beliefs and consumption, maximum betweenness centrality is negatively correlated with the observed bias. Golub and Jackson (2010) suggest that the influence of the most influential agent should be inversely related to convergence to the true state of the world. We, however, observe the opposite, when influence is measured in terms of betweenness centrality.

---

[53] A measure of how often each graph node appears on a shortest path between two nodes in the graph. $c = \sum_{s,t \neq u} \frac{n_{s,t}(u)}{N_{s,t}}$ is betweenness centrality of node u , where $n_{s,t}(u)$ is the number of shortest paths between s and t that pass through node u, and $N_{s,t}$ is the total number of shortest paths between s and t.

As network structure may play a critical role in accurately calibrating our control condition, we call into question, whether such baseline is meaningful to be constructed in the first place.

Figure 14: Consumption bias versus the maximum number of outward

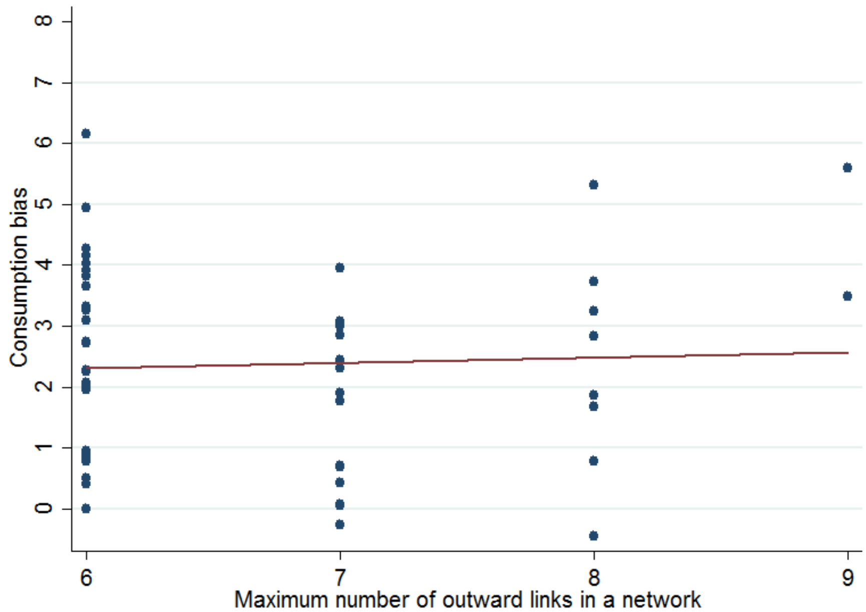


Figure 15: Belief bias versus the maximum number of outward links

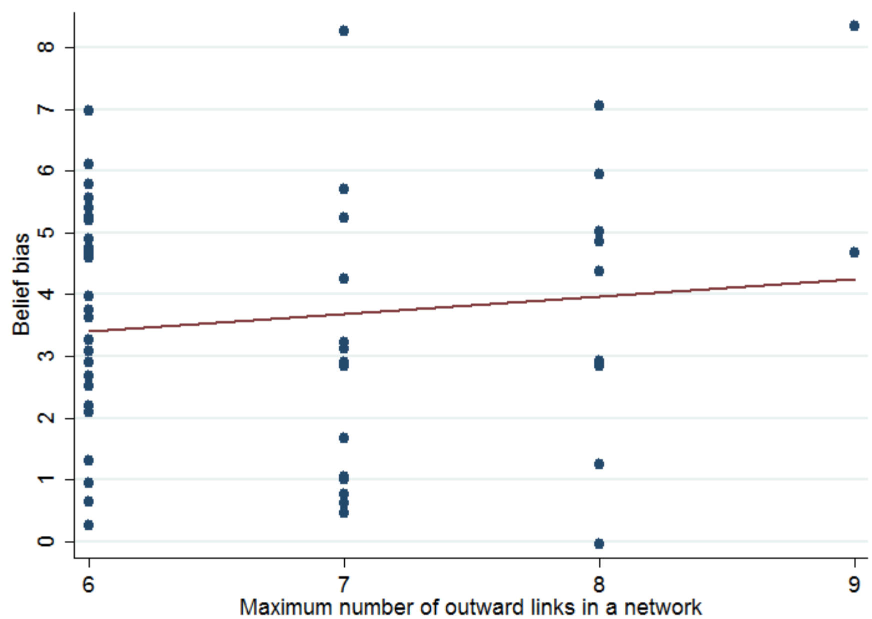


Figure 16: Consumption bias versus the maximum betweenness centrality score

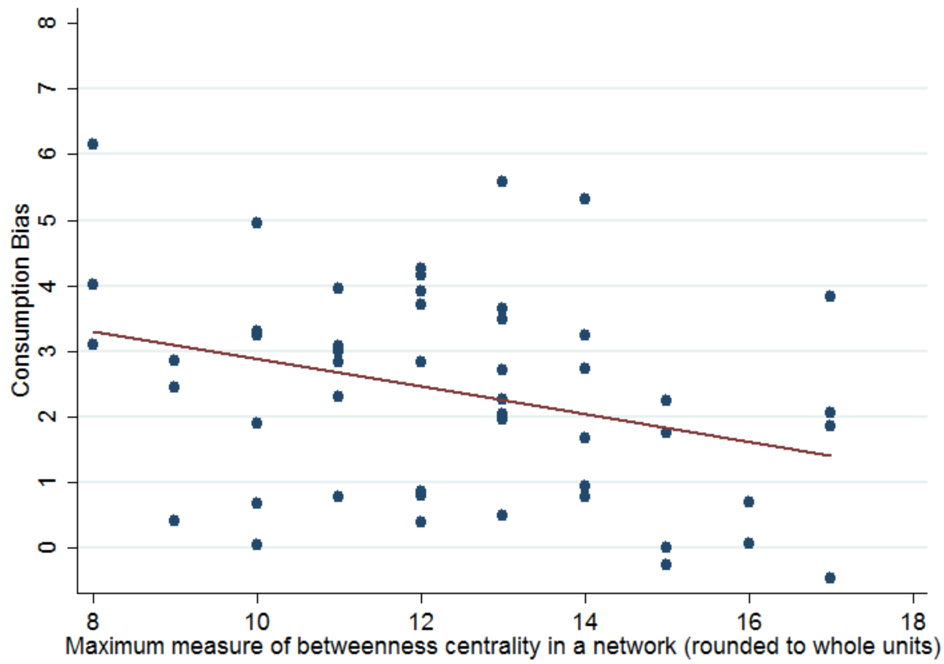
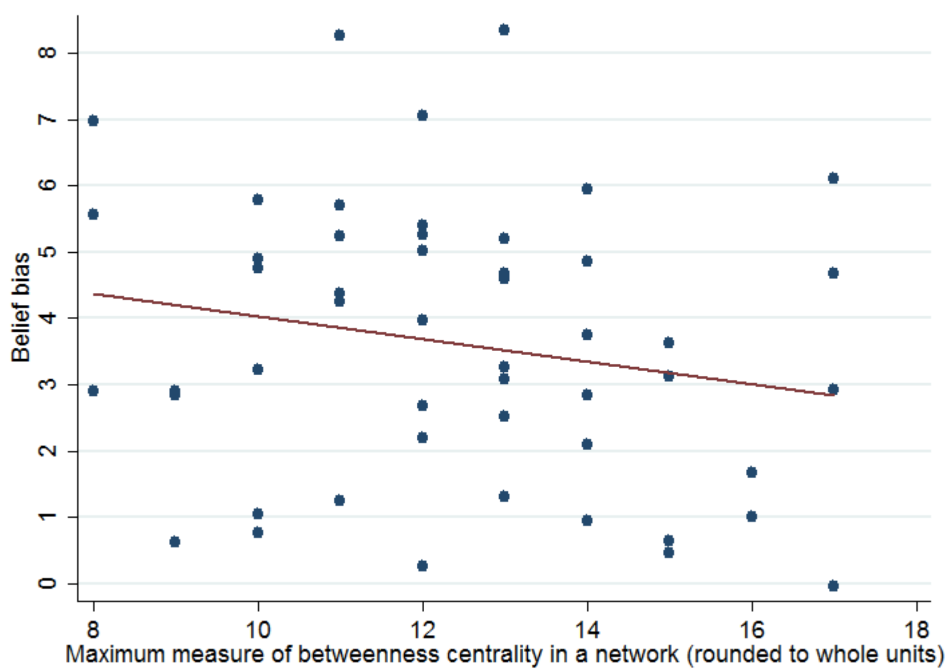


Figure 17: Belief bias versus the maximum betweenness centrality score



4.5 Discussion

The purpose of this chapter is to test, whether social norm misperceptions and overconsumption can be generated by the mechanism outlined in Jackson (2017). We construct a laboratory experiment, where we exogenously manipulate the structure of the imposed "social" networks. We observe a positive bias in the socially complementary consumption and perceptions about the average consumption of others in the treatment condition, where the correlation between the induced preferences and the number of individual network links is positive. However, bias also exists in the control condition, where the number of social links is fixed for all subjects. After controlling for belief-action consistency, we find that belief bias is significantly larger in the treatment condition. However, consumption bias is not significantly different between the treatment and control conditions.

We suggest multiple potential sources of the bias in the control condition. Rational, belief-action consistent, individuals seem to form beliefs, at least in part, as suggested by Jackson (2017): by anchoring their beliefs about the behavior of the group at the observed behavior of their social network neighbors. The belief-action inconsistent individuals, however, appear to select higher consumption levels than the rational players. Such behavior is sufficient to generate bias even in the control group. Moreover, some participants report to increase their consumption to manipulate the average consumption in the group, thus, such strategic behavior may exacerbate the bias. Hedging, however, does not seem to be a reasonable explanation for the observed bias as subjects exhibit a positive bias regardless of the state of the world they are assigned to.

Overall, we conclude that our experimental design, directly based on the theoretical framework in Jackson (2017), is not ideal to investigate the research questions we are interested in. In our future work, we plan to deviate from the abstract experimental setting presented here. Instead, we plan to follow the example of Kearns et al. (2009), who investigate biased voting in the context of social networks. Their Minority Power experiments already act as a great example of how minority preferences can override the majority preferences. To investigate Jackson (2017) mechanism, Kearns et al. (2009) design can be extended to represent consumption rather than voting decisions such that the coordination incentives are not of primary focus[54] and signals are not directly communicated, which may give rise to potential cursed equilibria (Eyster and Rabin, 2005).

Basic context should also be provided to the experimental participants such that we could focus on individual intuitive decision making rather than the ability to use and interpret the abstract experimental interface we employed in this chapter. Similarly, we would also employ multiple rounds rather than continuous decision making with feedback such that learning effects could take place. Subjects would also be paid for multiple randomly chosen rounds abandoning the exact functional payoff form from Jackson (2017), which would allow to increase payoff salience. A more intuitive context may allow us to avoid the issues associated with belief-action inconsistency.

Following the discussion of the literature on information aggregation in social networks, we acknowledge that a baseline, where norm misperceptions do not take place, is difficult to design and would not be very useful to study for our purposes as the zero bias result would heavily depend upon the properties of the social networks and social learning dynamics. Instead, we suggest it is useful to create a baseline, where norm misperceptions do occur.

[54] In Kearns et al. (2009) subjects earned nothing if consensus was not achieved.

Consequentially, we would like to investigate, how the bias varies with social network properties in multiple treatment conditions (for example, the existence of influential agents, network assortativity, homophily, clustering). Understanding of how network structure interacts with bias formation would allow us to identify the most problematic networks that require an intervention.

Primarily though, we are interested in investigating the effectiveness of an information intervention as a method to reduce social norm misperceptions and overconsumption. It is especially important to identify whether strategic incentives may prevent accurate information from spreading in the network. For example, most of the information interventions are currently implemented among the heavy alcohol users. If strategic incentives prevail such that heavy alcohol users manipulate the norm perceptions of their social acquaintances, information interventions should actually target individuals, who consume relatively low levels of alcohol.

To conclude, conducting this study has provided us with a greater understanding of belief formation in the social networks. We consider that the improved experimental design discussed above would create a useful platform for investigating biased norm perceptions, the abuse of the socially complementary behaviors, and the methods to battle both belief and behavior biases.

5. GENERAL CONCLUSION

In the three essays of this dissertation we investigate the role of the strategic beliefs in determining individual behavior. In particular, we discuss individual strategic beliefs in a simple strategic environment in Section 2, the effects of experimental design parameters on beliefs that may affect individual trusting behavior in Section 3, and the mechanism underlying biased strategic belief formation in the context of socially complementary behavior in Section 4.

In Section 2, we test whether the Cognitive Hierarchy model developed by Camerer et al. (2004) is more than simply an "as if" theory of behavior. Cognitive Hierarchy describes an individual as believing her opponents engage in heterogeneous steps of strategic thinking. To test for such beliefs, we run a two-phase experiment. In phase one, a group of subjects play games we design particularly to identify players of heterogeneous steps of strategic reasoning. In phase two, a different group of subjects predict their behavior. We find that a phase two participant anticipates a number of non-strategic individuals who play a naive strategy, players who play the best-response to this strategy, players who best respond to these strategic individuals and so on. We conclude that stepwise reasoning, the underlying assumption in the Cognitive Hierarchy model, provides insights on how individuals actually think in a strategic environment.

In Section 3, we discuss, how experimental design parameters may affect individual strategic beliefs, and, thus, measures of trust. This chapter replicates the original GLSS result: the survey and the incentivized measures of trust are not correlated, when a modified version of the original Berg et al. (1995) game is used. The modified game used in GLSS endows only the first mover, thus, changing the motives behind trusting behavior. The first mover in the modified game may be motivated by altruism or inequality aversion and, thus, may appear more trusting. The second mover, in turn, may return a smaller fraction of the amount received, in an attempt to equalize payoffs. This change in the second mover's behavior may, in turn, affect first mover's strategic beliefs, thus, distorting trusting behavior even further. In the reinvestigation experiment, we employ GLSS protocol; however, we replace the modified game with the original trust game design. We find that the survey measures of trust and behavior in the trust game are significantly correlated, when the standard game is used. We conclude that care should be taken in measuring complex economic behaviors, especially those, based on malleable strategic beliefs. Trust, for example, should be measured by the either responses to the survey questions or by the incentivized investment game, where both movers are equally endowed.

In the final Section 4, we set out to design a platform for investigating a potential underlying mechanism of social norm misperceptions (Jackson, 2017), and the methods to battle this belief bias as well as the associated over-engagement into complementary activity. We construct a laboratory experiment, where we exogenously manipulate the structure of the imposed "social" networks. We observe a positive bias in the socially complementary consumption and perceptions about the average consumption of others in the treatment condition, where the correlation between the induced preferences and the number of individual network links is positive. However, bias also exists in the control condition, where the number of social links is fixed for all subjects. After controlling for belief-action consistency, we find that belief bias is significantly larger in the treatment condition. However, consumption bias is not significantly different between the treatment and control conditions. We discuss the potential sources of bias in the control condition. We recognize that compliance with strict assumptions may be needed in the baseline to produce zero bias. We, thus, conclude that our experimental design is not ideal for investigating the research questions we are interested in. In our future work, we plan to deviate from the abstract experimental setting presented in Section 4. Instead, we plan to follow the example of Kearns et al. (2009) to construct a more intuitive, flexible and simpler design to answer multiple research questions of interest in our future work.

REFERENCES

Alaoui, Larbi and Penta, Antonio. 2016. Endogenous depth of reasoning. *The Review of Economic Studies*, 83(4):1297–1333.

Arad, Ayala and Rubinstein, Ariel. 2012. The 11–20 money request game: a level k reasoning study. *The American Economic Review*, 102(7):3561–3573, 2, 4, 22.

Armantier, Olivier, and Treich, Nicolas. 2013. Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, 62:17–40, 22.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. Trust, reciprocity, and social history. *Games and Economic Behavior* 10:122-142.

Bergstrom, R. L., Neighbors, C., and Lewis, M. A. 2004. Do men find "bony" women attractive?: Consequences of misperceiving opposite sex perceptions of attractive body image. *Body Image 1* : 183-191.

Bicchieri, C. 2010. Norms, preferences and conditional behavior. *Politics, Philosophy and Economics* 9(3) 297-313.

Bnabou, Roland and Jean Tirole. 2006. Incentives and Prosocial Behavior. *American Economic Review*, 96(5): 1652-1678.

Brown, A. L., Chua, Z. E., Camerer, C. F. 2009. Learning and Visceral Temptation in Dynamic Saving Experiments*. QJE.*

Brier, Glenn W. 1950. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3. 22

Brown, Alexander L, Camerer, Colin F, and Lovallo, Dan. 2013. Estimating structural models of equilibrium and cognitive hierarchy thinking in the field: The case of withheld movie critic reviews. *Management Science*, 59(3):733–747. 1

Camerer, Colin F. 2003. Behavioral game theory: Experiments in strategic interaction. *Princeton University Press*, 1.

Camerer, Colin F., Ho, Teck-Hua, and Chong, Juin-Kuan. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898. 1, 3, 5, 14, 15, 16, 23

Capra, C. Mónica, Kelli Lanier, and Shireen Meer. 2008. Attitudinal and behavioral measures of trust: a new comparison. Unpublished Paper, Emory University.

Carpenter, Jeffrey, Graham, Michael, and Wolf, Jesse. 2013. Cognitive ability and strategic sophistication. *Games and Economic Behavior*, 80:115–130. 22

Costa-Gomes, Miguel, Crawford, Vincent P., and Broseta, Bruno. 2001. Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5):1193–1235, 2.

Costa-Gomes, Miguel A., and Weizsacker, Georg. 2008. Stated beliefs and play in normal-form games. *The Review of Economic Studies*, 75(3):729–762, 22.

Ciriolo, Emanuele. 2007. Inequity aversion and trustees' reciprocity in the trust game. *European Journal of Political Economy* 23:1007-1024.

Crawford, Vincent and Costa-Gomes, Miguel. 2006. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768, 4.

Crosetto, Paolo, and Filippin, Antonio. 2013. The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, 47(1):31–65, 10.

Dominitz, Jeff and Hung, Angela A. 2009. Empirical models of discrete choice and belief updating in observational learning experiments. *Journal of Economic Behavior & Organization*, 69(2): 94–109, 22.

Dufwenberg, Martin and Gneezy, Uri. 2000. Measuring beliefs in an experimental lost wallet game. *Games and economic Behavior*, 30(2):163–182, 22.

Eister, Erik, and Rabin, Matthew. 2005. Cursed Equilibrium. *Econometrica,* 73(5): 1623-1672.

Elster, Jon. 1989. Social Norms and Economic Theory. *Journal of Economic Perspectives,* 3 (4): 99-117

Fehr, Ernst, Urs Fischbacher, Bernhard Von Rosenbladt, Jürgen Schupp, and Gert G. Wagner. 2003. A nation-wide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey. Working Paper, Institute for Empirical Research in Economics, University of Zurich.

Fischbacher, Urs. 2007. Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178, 7.

Fragiadakis, Daniel E., Knoepfle, Daniel T., and Niederle, Muriel. 2016. Who is strategic? 4.

Fragiadakis, Daniel E., Knoepfle, Daniel T., and Niederle, Muriel. 2017. Do individuals employ the same decision rules across strategic environments? 2, 21, 23.

Georganas, Sotiris, Healy, Paul J., and Weber, Roberto A. On the persistence of strategic sophistication. *Journal of Economic Theory*, 159:369–400, 2015. 2, 22, 23.

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman, and Christine L. Soutter. 2000. Measuring trust. *Quarterly Journal of Economics* 115: 811-846.

Goeree, Jacob K., Louis, Philippos, and Zhang, Jingjing. 2013.Noisy introspection in the “11–20” game. Technical report, Working paper. 22

Golub, Benjamin, and Jackson, Matthew O. 2010. Naïve learning in social networks and the wisdom of crowds. *American Economic Journal* 2: 112-49.

Good, Irving John. Rational decisions. 1952. *Journal of the Royal Statistical Society*. Series B (Methodological), pages 107–114, 22.

Greiner, Ben. 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association* 1: 114-125.

Jackson, Matthew O. 2017. The Friendship Paradox and Systematic Biases in Perceptions and Social Norms.

Johansson-Stenman, Olof, Minhaj Mahmud, and Peter Martinsson. 2013. Trust, trust games and stated trust: Evidence from rural Bangladesh. *Journal of Economic Behavior & Organization* 95: 286-298.

Johnson, Noel D., and Alexandra A. Mislin. 2011. Trust games: A meta-analysis. *Journal of Economic Psychology* 32: 865-889.

Johnson, Noel D., and Alexandra Mislin. 2012. How much should we trust the World Values Survey trust question?. *Economics Letters* 116: 210-212.

Kearns, Michael, Judd, Stephen, Tan, Jinsong, and Wortman, Jennifer. 2009. Behavioral Experiments on biased voting in networks. *PNAS*. 106 (5) 1347-1352.

Heap, Shaun Hargreaves, Arjona, David Rojo, and Sugden, Robert. 2014. How portable is level-0 behavior? A test of level k theory in games with non-neutral frames. *Econometrica*, 82(3): 1133–1151, 5.

Hortacu, Ali, Luco, Fernando, Puller, Steve, and Zhu, Dongni. 2017. Does strategic ability affect efficiency? Evidence from electricity markets. 1.

Huck, Steffen and Weizsacker, Georg. 2002. Do players correctly estimate what others do? Evidence of conservatism in beliefs. *Journal of Economic Behavior & Organization,* 47(1):71–85, 22, 23.

Kormos, Christine, Gifford, Robert and Brown, Erinn. 2015. The Influence of Descriptive Social Norm Information on Sustainable Transportation Behavior: A Field Experiment.

Environment and Behavior 47(5) : 479-501.

Larimer, Mary E. and Neighbors, Clayton. 2003. Normative Misperception and the Impact of Descriptive and Injunctive Norms on College Student Gambling. *Psychology of Addictive Behaviors*, 17(3) : 235-243.

Lazzarini, Sergio. G., Regina Madalozzo, Rinaldo Artes, and Jose de Oliveira Siqueira. 2005. Measuring trust: An experiment in Brazil. *Brazilian Journal of Applied Economics* 9: 153-169.

Nagel, Rosemarie. 1995. Unraveling in guessing games: An experimental study. The *American Economic Review*, 85(5):1313–1326, 3.

Palfrey, Thomas R. and Wang, Stephanie W. 2009. On eliciting beliefs in strategic games. *Journal of Economic Behavior & Organization*, 71(2):98–109, 22.

Perkins, H. W., Craig, David W. and Perkins, Jessica M. 2011. Using social norms to reduce bullying: A research intervention among adolescents in five middle schools. *Group Processes Intergroup Relations*,14(5) (2011) 703-722.

Perkins, H.W., Haines, M.P. and Rice, R. 2005. Misperceiving the College Drinking Norm and Related Problems: A Nationwide Study of Exposure to Prevention, Information, Perceived Norms and Student Alcohol Misuse. Journal of studies on alcohol: 470-477.

Roth, Alvin E. and Malouf, Michael W. 1979. Game-theoretic models and the role of information in bargaining. *Psychological review*, 86(6): 574, 2, 9, 20.

Selten, Reinhard. 1998. Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, 1(1):43–62. 22

Selten, Reinhard, Sadrieh, Abdolkarim and Abbink, Klaus. 1999. Money does not induce risk neutral behavior, but binary lotteries do even worse. *Theory and Decision*, 46(3):213–

252. 20.

Soroa-Koury, Sandra and Yang, Kenneth C.C. 2010. Factors affecting consumers' responses to mobile advertising from a social norm theoretical perspective". *Telematics and Informatics* 27 : 103-113.

Stahl, Dale O. and Wilson, Paul W. 1994. Experimental evidence on players' models of other players. *Journal of economic behavior & organization*, 25(3):309–327, 3.

Stahl, Dale O. and Wilson, Paul W. 1995. On players models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1, 5, 14.

Wang, Joseph Tao-yi, Spezio, Michael, and Camerer, Colin F. 2010. Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *The American Economic Review*, 100(3):984–1007, 2.

Wilson, Rick, and Catherine C. Eckel. 2011. Trust and Social Exchange. In *The Handbook of Experimental Political Science*, edited by James N. Druckman, Donald P. Green, James H. Kuklinski and Arthur Lupia. Boston: Cambridge, England: Cambridge University Press, pp. 243-257.

Xiao, Erte, and Cristina Bicchieri. 2010. When equality trumps reciprocity. *Journal of Economic Psychology* 31: 456-470.

Yan, Jin, and Lingling Miao. 2007. Effects of endowments on reciprocal behaviors. *2007 International Conference on Wireless Communications, Networking and Mobile Computing* (pp. 4591-4594), 21-25 September.

Young, Sean D. and Jordan, Alexander H. 2013. The Influence of Social Networking Photos on Social Norms and Sexual Health Behaviors. *Cyberpsychology*, Behavior and Social Networking, 16(4).

Young, Jacob T. N. and Weerman, Frank M. 2013. Delinquency as a Consequence of Misperception: Overestimation of Friends? Delinquent Behavior and Mechanisms of Social Influence. *Social Problems*, 60(3): 334-356.

APPENDIX 1

Table 18: Comparisons of subjects who only participated in the online survey with those who participated in both online survey and the experiment

	Survey Only	**Both Survey and Experiment**	**P-values**
GSS Trust	0.36 (0.49)	0.37 (0.49)	1.000
GSS Fair	0.51 (0.51)	0.54 (0.50)	0.868
GSS Helpful	0.55 (0.50)	0.52 (0.50)	0.739
Freshman	0.25 (0.44)	0.30 (0.46)	0.586
Male	0.46 (0.50)	0.49 (0.50)	0.743
White	0.40 (0.49)	0.48 (0.50)	0.407
Only child	0.40 (0.49)	0.43 (0.50)	0.741

Notes: One subject participated in the experiment without completing the survey. Our analysis does not include this subject and his partner.
Standard deviations are in parentheses. The p-values reported are computed by using Fisher's Exact Test.

Table 19: Original, replication and reinvestigation subject comparison

	GLSS	Replication & Reinvestigation	P-values
Male	0.33 (0.47)	0.49 (0.50)	0.004[a]
Freshman	0.69 (0.46)	0.30 (0.46)	0.000[a]
Only Child	0.11 (0.32)	0.23 (0.42)	0.005[a]
White	0.61 (0.49)	0.48 (0.50)	0.022[a]
GSS Trust	0.47 (0.50)	0.37 (0.48)	0.077[a]
Trust Index	0.01 (2.21)	0.00 (1.77)	0.000[b]

Standard deviations are in parentheses.

[a] Fisher Exact Test

[b] Kolmogorov Smirnov test

Table 20: Return ratio as a function of recipient characteristics

	GLSS					Replication			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Amount Sent	0.018*** (0.007)	0.018*** (0.006)	0.019*** (0.006)	0.014* (0.008)		0.013 (0.009)	0.014* (0.008)	0.015* (0.009)	0.006 (0.005)
Different sexes	0.003 (0.053)	-0.007 (0.052)	0.006 (0.055)	0.001 (0.065)		0.008 (0.067)	0.004 (0.067)	0.004 (0.071)	-0.023 (0.048)
Promise	-0.043 (0.051)	-0.007 (0.051)	-0.031 (0.052)	0.017 (0.063)		-0.051 (0.072)	-0.050 (0.070)	-0.067 (0.071)	0.078 (0.045)
Male	0.027 (0.059)	0.047 (0.058)	0.013 (0.061)	-0.015 (0.073)		-0.034 (0.075)	-0.032 (0.072)	-0.003 (0.071)	0.087 (0.049)
White	0.075 (0.054)	0.072 (0.052)	0.074 (0.055)	0.062 (0.065)		-0.026 (0.072)	-0.018 (0.072)	-0.045 (0.073)	0.004 (0.047)
Freshman	-0.072 (0.055)	-0.052 (0.055)	-0.083 (0.056)	-0.009 (0.071)		-0.035 (0.082)	-0.031 (0.079)	-0.071 (0.075)	0.014 (0.041)
Only child	-0.217** (0.092)	-0.242*** (0.089)	-0.218** (0.088)	-0.191* (0.112)		-0.072 (0.111)	-0.066 (0.110)	-0.076 (0.114)	-0.141** (0.062)
GSS trust	0.106** (0.051)					-0.081 (0.081)			
Trust index		0.043*** (0.012)					-0.020 (0.015)		
Self-reported trustworthiness			-0.026 (0.026)					-0.002 (0.038)	
Honesty index				0.01 (0.008)					0.004 (0.009)
Constant	0.174 (0.153)	0.165 (0.151)	0.373* (0.213)	0.261 (0.186)		0.453*** (0.128)	0.404*** (0.121)	0.422* (0.241)	0.354*** (0.066)
Adj.R^2	0.161	0.232	0.138	0.0359		-0.0276	-0.0051	-0.0573	0.2725
Observations	90	88	91	64		44	44	44	18

Note: The constant coefficients for GLSS are slightly different from the published version of the paper.
*Standard errors are in parentheses. Significance: * p<0.10, ** p<0.05, ***p<0.01*

Table 21: Return ratio as a function of recipient characteristics

	Reinvestigation			
	(1)	(2)	(3)	(4)
Amount Sent	0.016 (0.014)	0.019 (0.014)	0.024* (0.014)	-0.076 (0.037)
Different sexes	-0.138 (0.083)	-0.146 (0.086)	-0.170* (0.082)	1.050 (0.728)
Promise	-0.018 (0.072)	-0.001 (0.076)	-0.028 (0.072)	0.789 (0.526)
Male	-0.059 (0.076)	-0.041 (0.077)	-0.013 (0.076)	-0.619 (0.596)
White	-0.063 (0.08)	-0.068 (0.084)	-0.064 (0.079)	-0.090 (0.204)
Freshman	0.032 (0.113)	0.03 (0.118)	0.032 (0.111)	0.94 (0.786)
Only child	-0.12 (0.077)	-0.109 (0.079)	-0.06 (0.081)	-0.018 (0.240)
GSS trust	-0.1 (0.076)			
Trust index		-0.003 (0.018)		
Self-reported trustworthiness			0.095 (0.059)	
Honesty index				0.127 (0.495)
Constant	0.676*** (0.146)	0.595*** (0.14)	0.03 (0.372)	0.198 (0.495)
Adj. R^2	0.0604	-0.0072	0.00898	0.521
Observations	33	33	33	10

Note: The constant coefficients for GLSS are slightly different from the published version of the paper.
*Standard errors are in parentheses. Significance: * p<0.10, ** p<0.05, ***p<0.01*

APPENDIX 2

Table 22: Definitions of variables

Variable Name	Question/description	Answer range (Reinvestigation)
Amount sent	Amount sent by the first mover in trust game	0-10
Different Sex	First and second movers have different sex	Yes-1 No-0
Male	Subject is male	Yes-1 No-0
White	Subject is white	Yes-1 No-0
Freshman	Subject is freshman	Yes-1 No-0
Only Child	Subject is the only child in the family	Yes-1 No-0
GSS Trust	"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"	Most people can be trusted-1 Can't be too careful-0
Trust Index	Normalized index of responses to 3 trust questions: **1.** GSS Trust Question **2.** Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair? **3.** Would you say that most of the time people try to be helpful, or that they are mostly just looking for themselves?	(-3.6,2.5)
Return Ratio	Amount Returned/ (Amount Sent*3)	(0,1)
Self-Reported Trustworthiness	"I am always trustworthy"	Disagree Strongly-1 Agree Strongly-6
Honesty Index	Normalized index of response to four questions rating frequency of lying to parents, close friends, roommates, acquaintances partners.	(-10.8,6.2)

APPENDIX 3

A3.1 Key Questions from the Online Survey (Full Survey is available upon request).

This survey is designed to provide us with information about the study participants to help us gain additional insights into the results we will observe in the experiment. Your responses to this survey are guaranteed to be kept strictly anonymous. The survey should take approximately 20 minutes of your time. You will be paid $5 for completing the survey and $5 as a show up fee upon your arrival to the experiment. At the end of the experiment you will receive any earnings you will accumulate in the experiment. In order to get paid and participate in the experiment you must arrive to the Economics Research Laboratory on time. Please attempt to answer all the questions that follow to the best of your ability. Some questions are about facts (e.g. date of birth), and some of the questions are about your opinions. You are required to answer all the questions in order to participate in the experiment and receive your payment. This survey will take approximately 20 minutes. If for some reason you need to leave the survey, you can return using the link provided in the email. Any answers you have input into the survey will be saved. We appreciate your time. Thank you for your participation, The ERL Team

Q1. What is your sex?

❍ Female (1)
❍ Male (2)

Q2. What race do you consider yourself? ______________________

Q3. How many older brothers and sisters do you have? Include stepbrothers, stepsisters, and children adopted by your parents. (if your answer is zero, be sure to indicate zero)

Older Brother(s) (1)
Older Sister(s) (2)

Q4. How many younger brothers and sisters do you have? Include stepbrothers, stepsisters, and children adopted by your parents. (if your answer is zero, be sure to indicate zero)

Younger Brother(s) (1)
Younger Sister(s) (2)

Q5. What year are you in here at TAMU?

- ❍ Freshman (1)
- ❍ Sophomore (2)
- ❍ Junior (3)
- ❍ Senior (4)
- ❍ Graduate Student (5)

Q6 – Q9 and Q10b use the following response scale: Very often (1); Often (2); Sometimes (3); Rarely (4); Never (5); Prefer Not to Answer (6)

Q6. How often do you lie to your parents?

Q7. How often do you lie to your roommates?

Q8. How often do you lie to casual acquaintances?

Q9. How often do you lie to close friends?

Q10a. Do you have a steady boyfriend/girlfriend? Yes/No

If yes is selected:

Q10b: If you have a girlfriend/boyfriend, how often do you lie to her/him?

Q11. I am always trustworthy. (Indicate your level of agreement or disagreement with this statement.)

Disagree Strongly (1); Disagree Somewhat (2); Disagree Slightly (3); Agree Slightly (4); Agree Somewhat (5); Agree Strongly (6)

Q12. Would you say that most of the time people…

- ❍ try to be helpful. (1)
- ❍ are mostly just looking out for themselves. (2)

Q13. Do you think most people would try to…

- ❍ take advantage of you if they got a chance. (1)
- ❍ be fair. (2)

Q14. Generally speaking, would you say…

- ❍ that most people can be trusted. (1)
- ❍ you can't be too careful in dealing with people. (2)

A3.2 Script for the Experiment

Welcome, Please sit down anywhere in the lab with the person that you are paired with. You may ignore the stickers on the computer monitor for now.

Hi, thanks for coming to today's study.
I am [Experimenter], a graduate student here at Texas A&M University.

[Experiment Assistant] is our monitor today and will be helping with the experiment. We also have a monitor who will be helping behind the scenes of the experiment to make sure everything is completely confidential and she will be paying you in private at the end of the experiment.

We will be doing some activities and a survey because we are interested in how people make everyday decisions about money, and so in our studies people make some REAL decisions with REAL money.

I will be reading from a script today. We need to do that to make sure that we always say the same thing. From now until the end of the study, please do not talk to anyone except one of the experimenters unless instructed to do so. If you have any questions about the instructions please raise your hand.

You will receive $5 for coming here today. You will also receive $5 for completing the online survey. This $10 is yours to keep. There will be opportunities for you to earn additional money today: we will explain how as we go along.

Please listen carefully to all instructions that I give you. This is very important. The choice that you make will determine how much additional money you will be paid today.

When you came in today, you were given an ID card. You will later find a white card with identical ID number at the seat assigned to you. There are two unique colors of ID cards. Please take note which color you were assigned. We use ID numbers to keep track of the decisions that everyone makes, and you will need the number to claim your earnings at the end.

Social Connection Instructions:

Thank you for volunteering to participate in this economics experiment. The instructions you are reading are self-explanatory.

[Experiment Assistant] is going to pass out an envelope with two sets of instructions and the questionnaire inside now to each pair. Please make sure the Group ID number on the paper matches both of your personal ID numbers (the actual 2 digit number) on your colored card. Each of you can look at a copy of the instructions and follow along.

We will read these instructions aloud so please follow along. You and the person you are paired with in this experiment will always be reading the exact same documents; there is no deception involved. The instructions have been designed with an emphasis on clarity (which at times leads them to be a little repetitive). The experiment has been designed to use standardized procedures (which at times leads the experiment to feel a little formal). We will not answer any questions during this part of the experiment. If you have any questions, you should read back through the instructions. Now that the experiment has begun, please do not talk unless you are instructed to do so. However, if at any time you wish to end your participation in the experiment, please notify the experimenters immediately.

You have received either orange or green cards with your personal ID Number. Please remember whether you are ORANGE or GREEN in all future portions of this experiment.

ORANGEs and GREENs participate in most of the experiment in ORANGE-GREEN pairs. In this first portion of the experiment, which will take about ten minutes, each ORANGE-GREEN pair will get to know each other. This is the only exception to the experiment's NO-TALKING rule.

You will now spend ten minutes talking with each other and jointly filling out the enclosed ORANGE-GREEN QUESTIONNAIRE and the accompanying LIST PAGE. When you are finished, please be sure the total numbers of names on the List Page has been written down under Question 8 on the Questionnaire, and then have the ORANGE person pocket the List Page, so that researchers never see the names; this protects your anonymity. Put the completed ORANGE-GREEN QUESTIONNAIRE, into the WHITE ENVELOPE that it came to you in, and drop the envelope in the box indicated by the monitor. (Box is located at the front of the room) These instructions can then be trashed.

One way that might be easy to see who you know in common would be to look at your social media page (if you have one) and see if you have any friends in common. You may use your phone for this part of the experiment to see if there are mutual friends. Also using the survey will help you think about who you might know in common.

Begin talking for ten minutes when the monitor instructs you to do so. Please do not stop until the monitor instructs you to stop. Even if you do not think you know anyone in common, please keep trying to uncover some connections between the two of you, however remote they may be. It is important to use the entire 10 minutes.

Your ten minutes are over. Please place the survey back in the white envelope and bring it to the front of the room and place in the box. The orange member should hold on to the list page.

Now please find the computer with a sticker on the monitor that matches your id number and color. The green stickers are on my right side of the room and the orange stickers are on my left side of the room.

Also, please now turn off your cell phones and we will begin the next part of the experiment.

Transfer Game Instructions: (If Baseline – Promise treatment in parentheses)

Now you will receive a new decision form. The game we are about to play is a transfer game.

[Experiment Assistant] will now hand out an envelope with your group ID on the front and a decision sheet inside to the orange members of the pairs.

In this decision form: Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

There are two sections of this form. The top part will be filled out by the Orange player and the next portion will be filled out by the green player.

You and this other person will make decisions that affect both of your earnings. You both start with $10.

The orange player goes first, and can send none, some, or all of their $10 to the green player. Any money the orange player sends to the green player will be tripled.
The green player will have the $10 they started out with plus three times the amount of money sent to them by the orange player. The green player may send none, some, or all of the tripled money back to the orange player. Any money sent to the orange player will not be multiplied. Both decisions together determine the payoffs of both players.

Now Orange players make your decision to send some, none or all of your ten dollars to the green player. Orange players please make your decision now.
After making your transfer decision please place it back in the envelope and raise your hand. [Experiment Assistant] will come by and pick up your envelope, the outside experimenter will check for erroneous marks and then [Experiment Assistant] will redistribute the envelopes to the green members.
(Wait for envelopes to be picked up and redelivered to green)

[Experiment Assistant] has now returned the envelopes to the green players. Please open the envelope and you will see how much money was transferred to you by the orange player. Now please decide how much you would like to send back to the orange player. Remember this amount can be whole dollar increments. You can send some, none, or the entire tripled amount received.
Green players please make your decision now.

Raise your hand after you complete this decision and [Experiment Assistant] will pick the envelope with the decision form up. We will take the envelopes to the outside experimenter who will allocate the money for payment.

We will then begin a new task. This will be the bonus-winning task.

Transfer Game Instructions: (If Promise)

Now you will receive a new decision form. The game we are about to play is a transfer game.

[Experiment Assistant] will now hand out an envelope with your group ID on the front and a decision sheet inside to the green members of the pairs.

In this decision form: Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

There are three sections of this form. The first and third part will be filled out by the green player and the second portion will be filled out by the orange player.

You and this other person will make decisions that affect both of your earnings. You both start with $10.

The green player has an opportunity to make a non-binding promise to the orange player. They will do this by checking one of two statements discussed shortly.
After the green player makes the non-binding promise decision.
The orange player goes next, and can send none, some, or all of their $10 to the green player. Any money the orange player sends to the green player will be tripled.

The green player will now have the $10 they started out with plus three times the amount of money sent to them by the orange player. The green player may send none, some, or all of the tripled money back to the orange player. Any money sent to the orange player will not be multiplied. Both decisions together determine the payoffs of both players.

The promise that the green player can make is to repay the orange player at least as much as they sent. The other option is to make no promise. When the orange player receives the envelope the green player will either have made this non-binding promise by selecting the box indicating this. Otherwise, they will select the box indicating no promise. This promise will not affect the payments of either the orange or the green members.

After then green player completes this non-binding promise decision, please raise your hand, [Experiment Assistant] will pick up your envelopes and the outside experimenter will check for extraneous marks or invalid decisions. Then [Experiment Assistant] will bring the envelopes back and distribute them to the orange members of the pair for their transfer decision.

Green players please select if you would like to make a non-binding promise to the orange player now.

(Wait for envelopes to be picked up and redelivered to orange)

[Experiment Assistant] has now returned the envelopes to the orange players. Please open the envelope and you will see if the green player made a non-binding promise to return at least what you send to him.
Now Orange players make your decision to send some, none or all of your ten dollars to the green player. Orange players please make your decision now.

After making your transfer decision please place it back in the envelope and raise your hand. [Experiment Assistant] will come by and pick up you envelope, the outside experimenter will check for erroneous marks and then [Experiment Assistant] will redistribute the envelopes to the green members.
(Wait for envelopes to be picked up and redelivered to green)

[Experiment Assistant] has now returned the envelopes to the green players. Please open the envelope and you will see how much money was transferred to you by the orange player. Now please decide how much you would like to send back to the orange player. Remember this amount can be whole dollar increments. You can send some, none, or the entire tripled amount received.

Green players please make your decision now.

Raise your hand after you complete this decision and [Experiment Assistant] will pick the envelope with the decision form up. We will take the envelopes to the outside experimenter who will allocate the money for payment.

We will then begin a new task. This will be the bonus-winning task. [The instructions for the bonus-winning task is excluded from this script and it is available upon request.]

When you finish this please raise your hand, [Experiment Assistant] will pick up your bonus winning procedure decision sheets and we will give you a survey.

A3.3 Pre-Survey

INSTRUCTIONS **INDIVIDUAL ID:____________**
(one copy per person)

Thank you for volunteering to participate in this economics experiment. The instructions you are reading are self-explanatory. We will read these instructions aloud so please follow along. You and the person you are paired with in this experiment will always be reading the exact same documents; there is no deception involved. The instructions have been designed with an emphasis on clarity (which at times leads them to be a little repetitive). The experiment has been designed to use standardized procedures (which at times leads the experiment to feel a little formal). We will not answer any questions during the experiment. If you have any questions, you should read back through the instructions. Now that the experiment has begun, please do not talk unless you are instructed to do so. However, if at any time you wish to end your participation in the experiment, please notify the experimenters immediately.

You have received either orange or green cards with your personal ID Number, and you have an envelope that is either orange or green. Please <u>remember whether you are ORANGE or GREEN</u> in all future portions of this experiment.

ORANGEs and GREENs participate in most of the experiment in ORANGE-GREEN pairs. In this first portion of the experiment, which will take about ten minutes, each ORANGE-GREEN pair will get to know each other. This is the <u>only exception</u> to the experiment's NO-TALKING rule.

You will now spend ten minutes talking with each other and jointly filling out the enclosed ORANGE GREEN QUESTIONNAIRE (front and back) and the accompanying LIST PAGE. When you are finished, <u>please be sure the total numbers of names on the List Page has been written down under Question 8 on the Questionnaire (back)</u>, and then have the ORANGE person pocket the List Page, so that researchers never see the names; this protects your anonymity. Put the completed ORANGE-GREEN
QUESTIONNAIRE, into the WHITE ENVELOPE, and drop the envelope in the box indicated by the monitor. (Box will be located at the front of the room) These instructions can then be trashed.

Begin talking for ten minutes when the monitor instructs you to do so. Please do not stop until the monitor instructs you to stop. Even if you do not think you know anyone in common, please keep trying to uncover some connections between the two of you, however remote they may be. It is important to use the entire 10 minutes.

ORANGE -GREEN QUESTIONNAIRE **GROUP ID:_______________**

(one copy per pair)

Questions: (Try to come to agreement on the answers to these questions. If you absolutely cannot agree on a particular question, feel free to enter two different answers.)

1. Have the two of you ever met before? Yes/No
   1a. When did you first meet (month and year, MM/YYYY)?
   1b. Approximately how many times have you spoken in the last six months? ______
2. Do you live in the same house/apartment/dormitory? Yes/No
   2a. If yes: do you share a room? Yes/No
3. Are there any extracurricular activities in which you both participate? Yes/No
   3a. If yes, enter the number of extracurricular activities in which you both participate: __
4. Have the two of you ever spent time together socially (outside of formal extracurricular activities)? Yes/No
   4a. If yes: approximately how many times have you spent time together socially during the last six months? _______
5. Are the two of you in (or planning to be in) the same concentration? Yes/No
6. How would you describe your relationship?

_____ Know each other very well.

_____ Know each other pretty well.

_____ Know each other a little.

_____ Recognize each other.

_____ Don't recognize each other.

If the two of you know each other, continue with the lettered questions below; otherwise skip to question 7.

6a. Have you ever had any disagreements? (Pick one answer from the list below.)

_____ No, never.

_____ Only one minor disagreement.

_____ Occasional minor disagreements.

_____ One or two major disagreements.

_____ Many major disagreements.

6b. Would you describe your relationship as a friendship? Yes/No

6c. If yes: how close is your friendship?

_____ Extremely close friendship.

_____ Very close friendship.

_____ Close friendship.

_____ Casual friendship.

7. How likely is it that the two of you will see each other - intentionally or unintentionally - at some point in the next two weeks? (This includes accidental meetings, like seeing each other at MSC.)

_____ Very likely.

_____ Somewhat likely.

_____ Somewhat unlikely.

_____ Very unlikely.

8. Please use the List Page that came with this Questionnaire to list all of the people that you know personally and in common (a mutual acquaintance counts, Tiger Woods probably does not). Think about information like your house, concentration, extra-curricular activities, etc.., which will help you make this list as complete as possible). Finally, count all of the people on your list and enter that number here:_____.

Give the List Page to Person ORANGE, who should pocket it. We don't want the list because it might compromise your anonymity vis-a-vis the experimenters. Do not put any names from the List Page here on the Questionnaire.

LIST PAGE
(one copy per pair)

List all of the people that you know personally and in common (for example, a mutual acquaintance counts, but Tiger Woods does not). Remember to think about information like your house, concentration, extracurricular activities, etc., which will help you make this list as complete as possible.

When you are done, remember to count all of the entries on the list, and enter the number under question 8 on the ORANGE-GREEN QUESTIONNAIRE. Then give this page to Person ORANGE, who should pocket it. We don't want the actual names because they might compromise your anonymity vis-a-vis the experimenters.

A.3.4 Transfer Game Decision Sheet (No Promise, Reinvestigation)

DECISION FORM **GROUP ID______________**

Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

This section is to be completed by person **ORANGE**.

ORANGE's Transfer to GREEN
I, person ORANGE, choose to send $____("sent") of my $10 to GREEN (it must be a whole dollar amount) and choose to keep $____ ("kept"). I understand that GREEN is under no obligation to pay any of this money back. I also understand that any money that I give to GREEN will be tripled by the experimenter. So GREEN will actually receive three times what I send, or $____ ("tripled" - please do this calculation).

This section is to be completed by person **GREEN**.
Refer to the above amounts "kept" and "tripled", when you fill out this section.

GREEN's Transfer Back to ORANGE
I, person GREEN, choose to send $_____ ("sent back") back to ORANGE (it must be a whole dollar amount). I understand that I am under no obligation to send any money back to ORANGE. Now fill in the payout table below.

ORANGE gets	$ ("kept")	+	$ ("sent back")			, or	$	(Orange's winnings); and	
GREEN gets	$ ("tripled")	-	$ ("sent back")	+	$10	, or	$	(Green's winnings);	

Please leave the section below blank.
FOR USE BY CASHIER ONLY:

Confirmed that ____ ("kept") plus ____ ("sent") = $10; 3 times ____ ("sent") = ____ ("tripled"); and ORANGE's and GREEN's winnings above are calculated correctly and disbursed accordingly.

CASHIER:

A.3.5 Trust Game Decision Sheet (Promise, Reinvestigation)

DECISION FORM **GROUP ID______________**

Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

This section is to be completed by person GREEN.
One of these two statements must be checked and no other type of message is allowed. Please note that this promise is not binding. Check the box in front of the chosen option below.

GREEN's Message to ORANGE

□ I, person GREEN, promise to repay ORANGE at least as much as ORANGE sends me. For example, if ORANGE sends me $4, which will be tripled by the experimenters to $12, then I will repay ORANGE at least $4

□ I, person GREEN, make no promise to ORANGE.

This section is to be completed by person **ORANGE**.

ORANGE's Transfer to GREEN

I, person ORANGE, choose to send $____("sent") of my $10 to GREEN (it must be a whole dollar amount) and choose to keep $____ ("kept"). I understand that GREEN is under no obligation to pay any of this money back. I also understand that any money that I give to GREEN will be tripled by the experimenter. So GREEN will actually receive three times what I send, or $____ ("tripled" - please do this calculation).

This section is to be completed by person **GREEN**.
Refer to the above amounts "kept" and "tripled", when you fill out this section.

GREEN's Transfer Back to ORANGE

I, person GREEN, choose to send $_____ ("sent back") back to ORANGE (it must be a whole dollar amount). I understand that I am under no obligation to send any money back to ORANGE. Now fill in the payout table below.

ORANGE gets [$] + [$] , or [$] (Orange's winnings); and
("kept") ("sent back")

GREEN gets [$] - [$] + [$10] , or [$] (Green's winnings);
("tripled") ("sent back")

Please leave the section below blank.

FOR USE BY CASHIER ONLY:

Confirmed that ____ ("kept") plus ____ ("sent") = $10; 3 times ____ ("sent") = ____ ("tripled"); and ORANGE's and GREEN's winnings above are calculated correctly and disbursed accordingly.
CASHIER:

A.3.6 Transfer Game Decision Sheet (No Promise, Replication)

DECISION FORM **GROUP ID**______________

Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

This section is to be completed by person **ORANGE**.

ORANGE's Transfer to GREEN
I, person ORANGE, choose to send $____("sent") of my $15 to GREEN (it must be a whole dollar amount) and choose to keep $____ ("kept"). I understand that GREEN is under no obligation to pay any of this money back. I also understand that any money that I give to GREEN will be tripled by the experimenter. So GREEN will actually receive two times what I send, or $____ ("tripled" - please do this calculation).

This section is to be completed by person **GREEN**.
Refer to the above amounts "kept" and "tripled", when you fill out this section.

GREEN's Transfer Back to ORANGE
I, person GREEN, choose to send $_____ ("sent back") back to ORANGE (it must be a whole dollar amount). I understand that I am under no obligation to send any money back to ORANGE. Now fill in the payout table below.

ORANGE gets [$] + [$] , or [$] (Orange's winnings); and
("kept") ("sent back")

GREEN gets [$] - [$] , or [$] (Green's winnings);
("doubled") ("sent back")

Please leave the section below blank.
FOR USE BY CASHIER ONLY:

Confirmed that ____ ("kept") plus ____ ("sent") = $15; 2 times ____ ("sent") = ____ ("doubled"); and ORANGE's and GREEN's winnings above are calculated correctly and disbursed accordingly.

CASHIER:

A.3.7 Trust Game Decision Sheet (Promise, Reinvestigation)

DECISION FORM **GROUP ID______________**

Please write only what you have been asked to write. Extraneous marks, anywhere on this form, invalidate the Transfer Game; in that case neither one of you will be paid for your participation in this portion of the experiment.

This section is to be completed by person GREEN.
One of these two statements must be checked and no other type of message is allowed. Please note that this promise is not binding. Check the box in front of the chosen option below.

GREEN's Message to ORANGE

☐ I, person GREEN, promise to repay ORANGE at least as much as ORANGE sends me. For example, if ORANGE sends me $4, which will be doubled by the experimenters to $8, then I will repay ORANGE at least $4

☐ I, person GREEN, make no promise to ORANGE.

This section is to be completed by person **ORANGE**.

ORANGE's Transfer to GREEN

I, person ORANGE, choose to send $____("sent") of my $15 to GREEN (it must be a whole dollar amount) and choose to keep $____ ("kept"). I understand that GREEN is under no obligation to pay any of this money back. I also understand that any money that I give to GREEN will be doubled by the experimenter. So GREEN will actually receive two times what I send, or $____ ("doubled" - please do this calculation).

This section is to be completed by person **GREEN**.
Refer to the above amounts "kept" and "tripled", when you fill out this section.

GREEN's Transfer Back to ORANGE

I, person GREEN, choose to send $_____ ("sent back") back to ORANGE (it must be a whole dollar amount). I understand that I am under no obligation to send any money back to ORANGE. Now fill in the payout table below.

ORANGE gets	$	+	$	, or	$	(Orange's winnings); and
	("kept")		("sent back")			
GREEN gets	$	-	$	, or	$	(Green's winnings);
	("doubled")		("sent back")			

Please leave the section below blank.

FOR USE BY CASHIER ONLY:

Confirmed that ____ ("kept") plus ____ ("sent") = $15; 2 times ____ ("sent") = ____ ("doubled"); and ORANGE's and GREEN's winnings above are calculated correctly and disbursed accordingly.
CASHIER:

A.3.8 Post-Survey

TRANSFER GAME QUESTIONNAIRE INDIVIDUAL ID_______________

Please answer the following questions to the best of your ability. For most of the questions provide answers by circling a number on a 1 to 6 scale. Put a checkmark in front of the chosen multiple-choice and yes/no answers. Answer the open-ended questions to explain your opinions and experiences.

1. What do you think this portion of the experiment (The Transfer Game) was about?

____
____
____

2. How well did you understand the structure of the experiment?

(not at all well) 1 2 3 4 5 6 (very well)

Please explain.

____
____
____

3. How clear were the instructions written?

(clear) 1 2 3 4 5 6 (very clear)

Please explain.

____
____
____

4. How clear were the monitor's instructions?

(clear) 1 2 3 4 5 6 (very clear)

Please explain.

____
____
____

5. How good do you feel about your choices?

(good) 1 2 3 4 5 6 (very good)

Please explain.

____
____
____

6. In retrospect what, if anything, do you wish you had done differently? Please explain.

____
____
____

7. How much did you enjoy participating?
(not at all) 1 2 3 4 5 6 (very much)
Please explain.

8. Is there anything else you would like to tell us?

APPENDIX 4

Figure 18: Average consumption bias over time

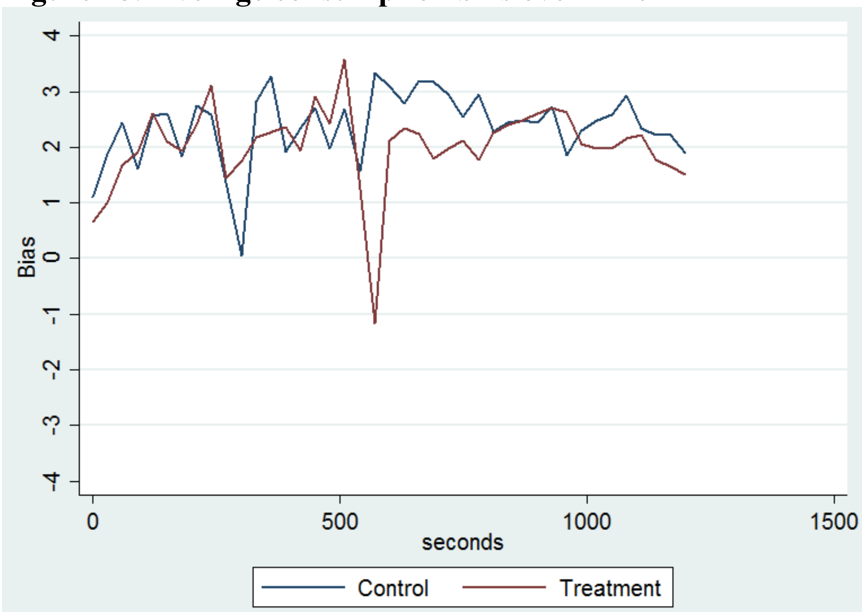


Figure 19: Average belief bias over time

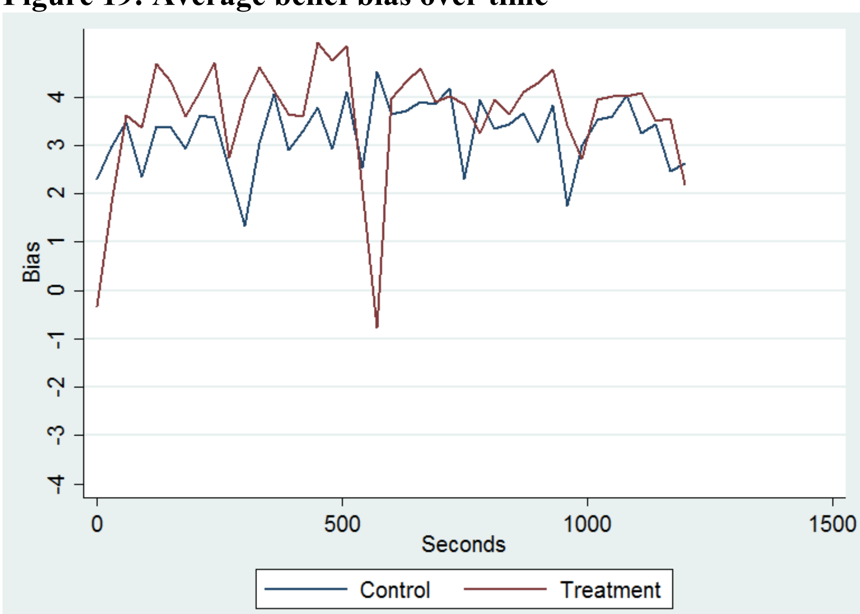


Figure 20: Average conditional consumption bias over time

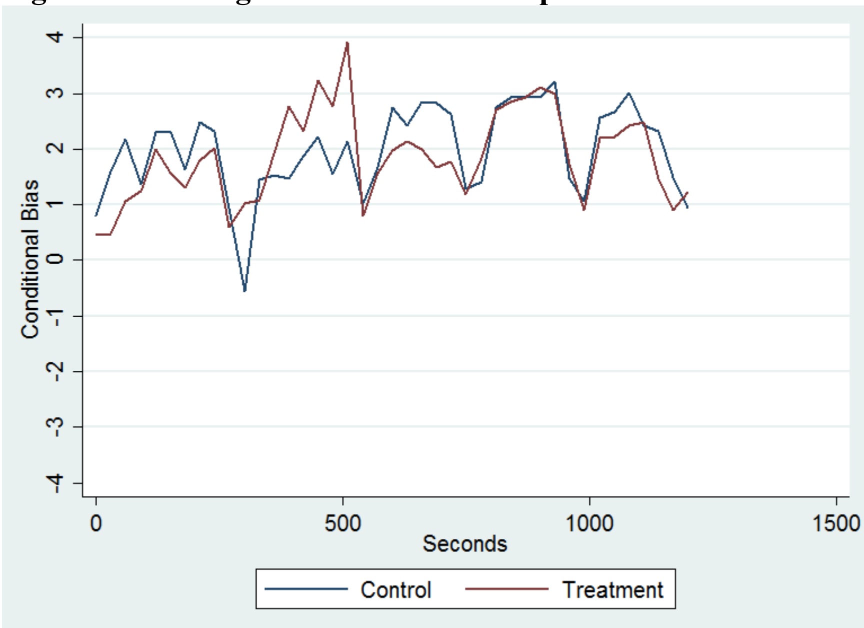


Figure 21: Average conditional belief bias over time