Essays on Information, Competition, and Cooperation

Thesis by

Luke Boosey

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended July 10, 2012)

Acknowledgements

After five wonderful years at Caltech, I am certain that these words will be not nearly enough to convey the full extent of my gratitude to all of my advisors, family, and friends. I have been fortunate to study under and alongside many of the brightest minds in the world, often without pause for the honor of such a blessing to sink in. I am truly thankful for all of the encouragement, advice, support, and understanding that I have been afforded during this period of my life.

To my advisor, John Ledyard, who has proved patient, supportive, and endlessly brilliant, I express my most sincere thanks. I am honored to count myself a student of his, certain as I am that there are few scholars from whom I could have learned so much. I will miss our discussions, be they about economics, beer, or baseball, and hope that I might continue to follow the path he has directed me to over the last few years.

I am also especially grateful to the other members of my committee — Federico Echenique, for introducing me to the theory of two-sided matching, and for quietly encouraging and developing my taste for economic theory; Leeat Yariv, who always asked the right questions and pushed me to explain and reason my intuition and ideas; and Jean-Laurent Rosenthal, for sharing with me his extensive knowledge about the world around us, and for guiding me (and others) through the last three years.

I could not go without acknowledging the tireless efforts of Laurel Auchampaugh, Tiffany Kim, Gail Nash, Susan Davis, Jenny Niese, and all of the administrative staff in the Division of the Humanities and Social Sciences. My life would have been far more difficult and far less enjoyable without their incredible dedication and expert assistance. I will be sad to say goodbye.

My education experience has been enriched by many people along the way, including my classmates, Salvatore Nunnari, Cary Frydman, Guilherme de Freitas, Sangmok Lee, Yutaka Kayaba, Vanessa Janowski, Muruvvet Buyukboyaci, Dustin Beckett, Julian Romero, Sera Linardi, Mohamed Mostagir, Ian Krajbich, my undergraduate advisor, Stuart McDonald, and Walter Yuan and Chris Crabbe, who made all the glitches disappear from my experimental software. My social experience at Caltech will be remembered fondly, due in large part to the great many friends I have come to know and love from all over the campus.

Finally, to my family, who are everything. To my sister, Amelia, and my brother, Thomas, who have been the greatest siblings anyone could ask for, even after I moved to the other side of the world. To my fiance and best friend, Kimberly, who has shown incredible love and support, and with whom I get to share the many upcoming adventures. To my mother, Beverley, who can lift my spirits in the blink of an eye, and whose unconditional love reminds me every day that I am blessed. And finally, to my father, Allan, who since the day I was born, has been my greatest supporter, my strongest advocate, my first (and best) advisor, my role model, and my hero; this is for you.

Abstract

This thesis consists of three papers that study the relationships between information, competition, and cooperation in two novel environments. We first examine the competitive behavior of firms with private information in two-sided matching markets. This part of the thesis employs purely game-theoretic tools. Second, we study voluntary contributions towards a linear public good by players who are connected through a network. In this environment, we use experimental and theoretical techniques to examine the effects of different information treatments and network structures on contributions.

In Chapter 2, I study the behavior of firms in a competitive market for workers. In particular, I study a game in which firms with private information compete for workers by committing to a single salary offer. Workers care only about salary and the matching process follows the deferred-acceptance approach introduced by Gale & Shapley (1962). For a two-firm, two-worker model, there exists a Bayesian-Nash equilibrium in which each firm type chooses a distributional strategy with interval support in the salary space. This equilibrium exhibits a *separation* of types, in the sense that types with a common most preferred worker choose nonoverlapping, adjacent supports. The type that makes higher offers is determined by the relative marginal value for the preferred worker. In larger markets, which replicate the two-firm, two-worker case, a comparable Bayesian-Nash equilibrium exists and the separation result endures. In the limit, there is no aggregate uncertainty about the realization of firm types, and competition is confined to the most popular worker type. Numerical results suggest that the finite market equilibrium strategies converge with replication to the corresponding equilibrium strategies in the limit case.

Chapter 3 studies individual contributions in a repeated public goods experiment. Subjects are

connected through a circle network, and consumption of the public good depends on a player's own contribution and the contributions of his neighbors. We study whether contributions depend on the nature of the information players are shown about others between rounds of the repeated game. We extend the approach of Arifovic & Ledyard (2009), which merges a modified model of other-regarding preferences (ORP) with a theory of learning. Our model predicts individual behavior that ranges from free-riding, to conditional cooperation, to unconditional giving. Many subjects switch between these different behavioral strategies across games with different information treatments. Individual contributions are remarkably consistent with our model, which combines other-regarding preferences, learning, and the information treatment. Both the data and model simulations suggest that learning (to play the benchmark Nash equilibrium of the game) is differential and contagious across players. Free-riders and unconditional givers learn faster than conditional cooperators, and provide an anchor that accelerates learning by their neighbors. These results suggest that the network or neighborhood structure may be important for contributions through its effects on learning.

In Chapter 4, we extend the analysis of learning and contributions in network public goods experiments. Using a set of five different network structures, we examine three key aspects of individual behavior. First, we report a negative finding regarding the predictions from our theory of other-regarding preferences. The theory provides certain predictions about how a particular subject should and should not behave across networks. We find several violations of these predictions, particularly in small, complete network groups, but also in the larger, more interesting networks. Second, we report on the average contributions by players in groups that consist of all conditional cooperators. In the one-shot game, these groups have a continuum of equilibria, in which every player contributes the same amount. While one might expect contributions to average half of the endowment, we find in both the data and learning simulations that average contributions decline over time to less than half of the endowment. We conjecture that learning dynamics may provide a method of equilibrium selection, for players trying to coordinate on one equilibrium in the repeated game.

Our main finding in this chapter is that learning is contagious in networks other than the circle, which we studied in Chapter 3. We find considerable evidence at the individual match level that free-riders and altruists provide an anchor that stabilizes behavior and accelerates learning by their conditional cooperator neighbors. Our analysis highlights the possibility that, even when the distribution of free-riders, altruists, and conditional cooperators is the same across networks, the different neighborhood structures may affect contributions differently through their effects on learning. Thus, the main contribution of this chapter is the confirmation that learning is contagious across a range of different network structures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The three papers in this thesis share a common focus on strategic settings where information plays a prominent role. We study two novel settings in which information is particularly important; (1) salary competition in two-sided matching markets when there is private information; and (2) voluntary contributions in a network public goods game.

The first setting is studied under complete information by several prominent authors, including Bulow and Levin (2006), Kojima (2007), and Niederle (2007). However, there is relatively little work that examines the problem when there is private information. In Chapter 2, we study the equilibrium salary offers made by firms with private information, who compete for workers in a two-sided matching market.

In Chapters 3 and 4, we switch focus to voluntary contributions in network public goods games. This setting is part of a newer, developing literature that studies behavior in games played on networks.[1] The first study of public goods on networks is Bramoullé and Kranton (2007), who prove for any social network structure, the existence of a Nash equilibrium in which self-interested agents either contribute everything or nothing. In this thesis, we consider different informational assumptions regarding what the agents know, or can observe about the network. Moreover, we assume that agents can have other-regarding preferences, rather than make the classical assumption that agents are purely self-interested.

In the experimental literature on voluntary contributions to linear public goods, there is a well-

[1]A fairly comprehensive study is conducted by Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010) for an extensive range of games.

documented inconsistency between the theory and experimental data. This has generated a great deal of work on alternative theories, including other-regarding preferences, repeated game strategies, and learning. Although these have met with some success, none of the alternative theories has provided on its own, a successful explanation for all the stylized facts that are evident in the experimental data. One of the more successful approaches is the recent theory provided by Arifovic and Ledyard (2009), who merge a model of other-regarding preferences with an evolutionary-style learning model called Individual Evolutionary Learning (IEL). Arifovic & Ledyard use their model to produce simulated behavior that is remarkably close to the data observed across a wide range of public goods experiments in the existing literature. We provide a full behavioral model that extends the approach developed by Arifovic and Ledyard (2009). We extend the model to the network public goods environment, and incorporate the different informational assumptions into the learning dynamics.

One of the key findings in the experimental public goods literature is that many players are *conditional cooperators*, in the sense that they tend to contribute more when others contribute more (and less when others contribute less). This notion of conditional cooperation raises two motivating questions, which we consider in Chapter 3. First, *how does conditional cooperation depend on the nature of information players observe about others between rounds of a repeated game?* We argue that information can inform both how, and on what the players condition their contribution decisions. Second, *how does conditional cooperation depend on the structure of interaction between players?* In particular, since players are connected through a network, information about a player's neighbors' neighbors may be important for determining how much to contribute.

The main contribution in Chapters 3 and 4 is the finding that learning (to play the Nash equilibrium of the network public goods game) is contagious across several different network structures. This result suggests an important relationship between learning and neighborhood structure in games played on networks.

In Chapter 3, we focus on a single network structure (the circle network), while varying the information players are shown between rounds of a repeated network public goods game. We first show that variation in the informational treatments within the experiments can induce a subject with stable underlying preference parameters to switch the form of his equilibrium behavior between free-riding (always contributing zero), unconditional giving (always contributing his endowment), and conditional cooperation. The data exhibits a significant amount of switching within-subject, which we show can be attributed to changes in the informational treatment, rather than a product of instability in players' underlying types. We then show that, while contributions typically do not converge to the Nash equilibrium of the static version of the game, they are consistent with the simulated behavior generated by our full behavioral model. In this chapter is where we first provide evidence for the main result that learning is contagious throughout the network.

In Chapter 4, we extend the analysis to a wider set of network structures. In all of the network environments, players participate in a 10-period repeated network public goods game. After each round, subjects observe the average payoff in their neighborhood, which is defined by network. As in Chapter 3, our specification of other-regarding preferences provides a set of cutoff conditions on a player's preference parameters, which allow us to classify players as free-riders, conditional cooperators, or pure altruists in any given game. From a static point of view, the classifications within-subject are inconsistent with the predictions of our theory of preferences, although these may be exacerbated by sample size issues or the coarseness of our classification criteria. Nevertheless, this negative result merely highlights further that contributions across different networks are sensitive not only to the composition of free-riders, altruists, and conditional cooperators, but also to the configuration of those players on the networks.

Chapter 4 also analyzes average contributions in groups comprising all conditional cooperators, which according to the static other-regarding preferences model, possess a continuum of equilibria. We find that both the data and learning simulations for these environments decline steadily over time, to an average contribution below half of the endowment. We conjecture that this decline is related to the learning dynamics, and that the learning process may provide an interesting method for equilibrium selection in these groups with multiple Nash equilibria.

Finally, our main finding in Chapter 4 is that learning is also contagious across the additional network structures, particularly in the connected cliques and two-step circle networks. We study the data at the individual match level and compare the results with some simple predictions of contagious learning in each realized match. For the most part, the data is consistent with the learning simulations, but even when it is not, the contributions exhibit a pattern consistent with contagious learning. Thus, our primary contribution from Chapters 3 and 4 is to show that contributions depend on the network structure through its effects on learning by conditional cooperators.

Chapter 2

Salary Competition in Two-Sided Matching Markets with Private Information

2.1 Introduction

In many labor markets, firms compete with each other for workers along several dimensions. These include salary, employee benefits, bonuses, health insurance coverage, and opportunities for career advancement. Firms often set policies regarding benefits, bonuses, health plans, and vacation time, rather than personalize the terms of employment for each individual worker. It is also common for firms to decide on a salary for a particular position, rather than negotiate a salary with each individual employee. These terms of employment can be inflexibile, either because they are firm-wide policies, contractual obligations, or because the salary for the position has been widely advertised. There are other settings in which agents on one side of a market make a costly investment in order to compete for the services (or affections) of the agents on the other side of the market. However, this paper focuses on the case of salary competition between firms. When choosing the terms of employment to offer to workers, each firm considers the preferences (or types) of the other firms. However, in most cases, the firms do not have full information about each other. This paper examines the competitive behavior of firms when they do not know each others' preferences.

There are several other papers that study salary competition in two-sided markets with complete information, in particular by Bulow and Levin (2006), Niederle (2007), and Kojima (2007).[1] Bulow and Levin (2006) consider the effect of a centralized matching mechanism on salary levels, and argue that the National Resident Matching Program (NRMP) algorithm compresses and depresses the salaries of workers, relative to the competitive equilibrium. These papers are also related to earlier research that deals with matching firms to workers by incorporating salary offers. The earliest treatment of matching with salaries appears to be Shapley and Shubik (1972), which modifies Gale and Shapley (1962) to incorporate a transferrable utility good in which salaries can be paid. The literature was further developed by Crawford and Knoer (1981) and subsequently, by Kelso and Crawford (1982), who devised a salary adjustment process which converges to a core allocation.[2] More recently, Hatfield and Milgrom (2005) develop a model of matching with contracts that incorporates the Kelso and Crawford (1982) model.[3] They show that if the preferences of the firms satisfy a *gross substitutes* condition and a *law of aggregate demand* condition, then truthful reporting is a dominant strategy for workers in a worker-proposing matching mechanism.

While matching with salaries has attracted considerable interest, the existing literature does not include a model in which the firms' preference orderings over workers are private information. Hoppe, Moldovanu, and Sela (2009) introduce a model of assortative matching in which there is incomplete information on both sides of the market, however, the incomplete information in their setting relates to attributes of potential partners, rather than preference orderings of potential competitors. Once a worker has chosen a signal in their model, every firm has the same ranking over workers, based on their signals. Likewise, once a firm has chosen a signal, every worker has the same ranking over firms. In contrast, the objective of this paper is to understand the effects of private information about the preference orderings of potential competitors on salary competition. The model extends the approach taken by most of the literature by allowing firms to have different primitive preference orderings over the set of workers. This assumption then allows for firms' preferences to be private

[1] Another paper that investigates the importance of various assumptions in the Bulow and Levin (2006) model is Gonzalez-Diaz and Siegel (Forthcoming), who focus on a set of job features, including salaries, reputation, responsibility, work hours, training, and quality of facilities, that may affect a hospital's attractiveness to workers in a nonlinear manner.

[2] A core allocation in this context is a one-to-one matching along with a salary schedule, in which no firm and no worker can negotiate a salary at which they would prefer each other over their current partners at their current salaries.

[3] Other related work includes Hoppe, Moldovanu, and Sela (2009), and Crawford (2008).

information, which is the main point of departure from the existing literature on matching with salaries.

Outline

Consider a game in which firms with private information compete for workers by making a single salary offer. Suppose each firm wants to hire at most one worker, and each worker can work for only one firm. Firms have types, defined by their preferences over the set of workers, and the option to remain unmatched. A key assumption is that firms' types are private information, although there is a common prior distribution over the type space. On the other hand, the workers all rank the firms in order of salary, from highest to lowest.[4] A firm must commit to a single salary offer, which they are required to pay to the worker with whom they are eventually matched. If they are unmatched, they do not pay anything. Although the problem is presented as one of choosing salaries to offer to potential workers, the firms' decisions can also be interpreted as investments in other terms of employment (such as benefits programs, health plans, or available facilities) that make the firm attractive to potential employees. Under either interpretation, firms are denied the flexibility to personalize the offers made to different workers for the same position.

Once salaries are chosen, the firms and workers are matched in the following manner. Each firm makes an offer to at most one worker. Each worker tentatively accepts at most one offer and rejects all the other offers it receives. Any firm who is rejected then makes the same offer to another worker who has not already rejected the firm. When no new offers are made, all remaining tentative matches are confirmed. This matching process is analogous to the firm-proposing deferred acceptance algorithm introduced by Gale and Shapley (1962) to prove the existence of stable matchings. A nice property of the firm-proposing deferred acceptance algorithm is that it gives firms a dominant strategy to make offers in a straightforward manner. In the present environment, this means that firms have a dominant strategy to make offers in order of preference, but only to workers who are acceptable to the firm at its chosen salary. Furthermore, since all workers care only about salary, there is a unique stable matching for any profile of firm preferences and salary offers. As a result, no worker has an

[4]This is a significant assumption, which eliminates the possibility of strategic behavior at the matching stage, as it induces a unique stable matching.

incentive to strategically reject an offer. Thus, the focus in this paper is on the salary decisions made by the firms.

In this paper, we begin by analyzing a two-firm, two-worker model, in which the firms can be one of four types. Type a prefers worker w_1 to worker w_2 whilst type b considers only worker w_1 to be acceptable. Similarly, type c prefers w_2 to w_1 while type d only considers w_2 to be acceptable. In the two-firm, two-worker model, there are no pure strategy equilibria. However, there exists a Bayesian Nash equilibrium in mixed (distributional) strategies which are continuous with interval support.

The first result of the paper is that the equilibrium exhibits a *s*eparation of types, in the sense that between two types who have a common most preferred worker, one type always makes higher offers than the other type. For example, given certain parameters, every salary offered with positive probability by a type b firm (who only wants worker w_1) is higher than any salary offered with positive probability by a type a firm (who also wants w_1 the most, but considers w_2 to be acceptable). That is, all type a firms "accept defeat" in case the other firm is a type b firm. Instead, type a firms concentrate on competing against another type a firm. The relative marginal value attached to the workers by different types determines which type makes the higher offers in equilibrium. When the realized types do not share a common most preferred type, there is no competitive pressure on the salary offers. Therefore, the relative probabilities of being either type a or type b, compared with type c or type d, also affects the equilibrium salaries. Furthermore, more popular workers (as determined ex-ante by the distribution over firm types), attract higher average equilibrium salaries. Likewise, as one might expect, the higher the probability of facing a given firm type, the higher the average salary offered by that type of firm in equilibrium.

The paper extends the analysis to larger markets by replicating the two-firm, two-worker model. The second result proves the existence of a Bayesian Nash equilibrium in continuous distributional strategies with interval support for each finite replicated market. The proof, which is by construction, also establishes the *s*eparation result for types with a common most preferred worker class. In the limit, when there are a continuum of firms and a continuum of workers, there is no aggregate uncertainty about the realization of types. Thus, competition in equilibrium is confined to the most popular worker class. Finally, numerical simulations suggest that the finite market equilibrium strategies converge to the corresponding continuum equilibrium strategies as the number of replications approaches infinity.

A related environment with a fixed, or posted wage is studied by Burdett & Mortensen (1998). They study a game where a continuum of firms choose permanent wage offers and a continuum of workers search by sequentially sampling from the set of offers. Workers search both while unemployed and while employed for a job with an acceptable, or higher wage, respectively. Firms post a wage conditional on the search behavior of the workers and the wages offered by other firms.

The principal result in Burdett & Mortensen (1998) is that wage dispersion is a robust outcome when workers must search for individual offers, provided that workers search while employed as well as when unemployed. They characterize the unique equilibrium (steady state) distribution of wage offers under different assumptions about firm and worker heterogeneity. The equilibrium distribution is non degenerate, exhibiting wage dispersion, even when all firms and workers are respectively identical, *as long as* the arrival rate of job offers is strictly positive, but finite, for all workers (even those already employed).

There are several important differences between the setup and the matching process in Burdett & Mortensen (1998) and our approach in this paper. First, in this paper, firms have single-unit demand. That is, each firm wants to be matched with just one worker, rather than build up a team of workers. Second, workers are not identical in our model. Offers are directed by firms to particular workers, rather than posted for workers to search for and accept as they please. This reflects the assumption that firm productivities depend also on which worker they employ and the additional heterogeneity in firms' ordinal preferences. In particular, some workers are not acceptable to a firm at that firm's posted salary.

An important implication of these differences is that, in our model, each worker faces a potentially different distribution of offers. Firms have both different ordinal preferences (and productivities) and may offer different salaries, whilst controlling which workers may receive their offer. Nevertheless, the models are related, since we can interpret the 'continued search' feature of Burdett & Mortensen (1998) as an ongoing matching process which evolves towards a stable, steady state equilibrium matching. In this respect, the assumption that employed workers continue to search for higher paying jobs is equivalent to a provisionally matched worker receiving a higher salary offer in a deferred acceptance mechanism.

In our paper, the equilibrium salary distributions offered by different firm types also exhibit wage dispersion (both within and among the types). However, the wage dispersion is driven by private information among the firms and the competition by heterogeneous firms for heterogeneous workers. In contrast, the wage dispersion in Burdett & Mortensen (1998) is driven in part by the multi-unit demands of the firms, and their heterogeneous productivities with respect to a set of perfectly substitutable workers. In fact, in this paper, the highest productivity firm for a given worker is not necessarily matched with that worker. Workers are instead sorted to firms based (generally) on the firms' relative *marginal* values for the different workers, which depend on the firms' ordinal preferences as well as the relative intensity of preferences (marginal productivities) for each of the workers, compared with the other firms.

Another related paper is Bulow and Levin (2006). In their paper, Bulow and Levin provide a concise argument that in a two-sided market with a centralized matching procedure, equilibrium salaries are lower than the corresponding competitive equilibrium. The matching of firms to workers is still highly efficient, which implies that the lower salaries are mostly offset by higher profits for the firms. Moreover, they show that the equilibrium supports for the salary distributions offered are more compressed for better workers, such that the most productive firms gain the most from the centralized match.

Both the setup and the results of our paper can be distinguished from Bulow and Levin (2006). The primary innovation in our model is to allow different firms to have different *ordinal* preferences over the same set of workers. This extends the approach followed in almost all the previous literature, including Bulow and Levin (2006). Also notice that this innovation is not the same as allowing for heterogeneous productivity types if the heterogeneity does not lead to a different preference ordering.

Rather, it allows for a broader range of types, and facilitates the introduction of a richer notion of private information to two-sided matching markets.

In Bulow and Levin (2006), each firm brings a degree of productivity to their match. Likewise, each worker possesses a fixed level of productivity. The value created in any match is determined by multiplying together the two productivities. One implication of this setup is that each of the set of firms, and the set of workers can be objectively ranked or organized from most productive to least productive. Then the efficient matching outcome involves an assortative match between the most productive firm and worker, followed by the second most productive firm and worker, and so on. In their model, productivities are complete information, and more importantly, every firm has the same preference ordering over the set of workers.

In this environment, there is no scope for introducing private information, except about the firm's own ranking among the other firms. But regardless of where the firm ranks in terms of productivity, all firms will share the same preference ordering over workers, such that the problem becomes a matter of determining 'where you stand' in the objective ranking of the firms. In contrast, the features of our model create uncertainty about the productivities of the agents on the other side of the market. That is, firms do not all agree on the value of matching with a particular worker. This gives rise to a situation in which, even though one firm may be less productive than another, the latter firm values the former firm's second choice worker more than his first choice worker. By removing the restriction that all firms prefer the same workers, we create situations where the degree of competition between firms is uncertain.

In complete information environments, firms can perfectly deduce the level of competitive pressure they will face when attempting to hire a particular worker. For instance, in the Bulow and Levin (2006) setup, firms observe the productivities of other firms and can predict which firms they will need to compete with or outbid to be matched with a particular worker. In our setup, there is an additional dimension which affects the level of competitive pressure. Depending on the realization of firms' types, a competitive equilibrium may distribute all the surplus from matches to the firms (salaries close to zero), split the surplus between workers and firms, or even transfer all of the surplus to the workers. For instance, in a two-firm, two-worker market, when both firms want worker a and only worker a, the strong competitive pressure will drive the salary up to the lower of the two firms' marginal values for worker a. On the other hand, if one firm prefers worker a to worker b, while the other prefers worker b to worker a, there is no competitive pressure, and the firms can post a salary equal to (or close to) zero, retaining all surplus from the match.

Furthermore, in this setting, even with complete information, the centralized matching equilibrium salaries do not always depress salaries relative to the corresponding competitive equilibria. When firms have different most preferred workers, the matching equilibrium salaries are equal to the lowest competitive equilibrium salaries, but when firms have the same most preferred worker, the matching equilibrium salaries equal the highest corresponding competitive equilibrium salaries.

Our results reflect that with private information, the firms face an uncertain level of competitive pressure. Thus, the equilibrium salaries tend to redistribute the surplus more evenly between firms and workers. Since individual firms do not know the precise realization of all firms' types, they are unable to extract all surplus when competitive pressure is low, nor are they forced to transfer all surplus to workers when competitive pressure is high.

We also show that competition between firms is localized. Specifically, when two different firm types have a common most preferred worker, each type focuses on competing against other firms of their own type. For example, if one type prefers a to b to remaining unmatched, while the other type also prefers a most but prefers to be unmatched over b, the type with a higher marginal value for a will pay a 'premium' large enough to outbid a firm of the other type, then choose a salary distribution that focuses on competing with another firm of their own type. This separation result is the primary feature of the equilibrium characterization in this paper, for both the two-firm, two-worker markets, and the larger, replicated markets. This localized nature of competition is different from the notion of localized competition discussed by Bulow and Levin (2006). In their setup, localized competition occurs in a framework where competitive pressure is readily observed, so that firms know which worker they are likely to be matched with, and which other firms (those with similar productivity) will be competing with them for that worker.

The rest of the paper is organized as follows. Section 2.2 describes the model and the matching process, then introduces the general two-firm, two-worker game with private information. A simple example demonstrates the main features of the equilibrium and the intuition behind the behavior of the firms. Section 2.3 characterizes the symmetric Bayesian Nash equilibria for which strategies are continuous distributions with interval support. The characterization also establishes the *s*eparation result for a general two-firm, two-worker model with private information. In Section 2.4, the model is extended to larger markets. This section characterizes the equilibria of the limit case in which there are a continuum of firms and workers, then proves existence for finite replicated markets. Finally, Section 2.5 presents the numerical simulations, which show that the equilibrium strategies in finite replicated markets converge to the corresponding continuum equilibrium strategies as the number of replications approaches infinity.

2.2 A Two-Firm, Two-Worker Model

Suppose there are two firms $f_1, f_2 \in F$ and two workers $w_1, w_2 \in W$. Each firm has strict preferences over the set $\{w_1, w_2, \emptyset\}$, where $\emptyset$ represents being unmatched. It is safe to ignore any preference ranking in which remaining single is the most preferred option. Thus, there are four possible preference rankings for each firm.

$$P_a : w_1\, w_2\, \emptyset \qquad P_b : w_1\, \emptyset\, w_2$$

$$P_c : w_2\, w_1\, \emptyset \qquad P_d : w_2\, \emptyset\, w_1.$$

Assume that each preference ranking is represented by a pair of values, one for each worker, while the value of remaining unmatched is 0. This assumption is somewhat restrictive, since it means that two firms with same preference ranking also have the same values for the workers. In Section 2.6, I discuss ways to relax this assumption about the type space.

Refer to a firm with preferences P_k as a firm of type k. Then the set of firm types is described as $\mathcal{P}^f = \{a, b, c, d\}$ where, for example, $a = (a_1, a_2)$ and a_j is the value of worker j to type a for each $j = 1, 2$. In order to represent the preference rankings, the values of the different types must satisfy the following conditions.

$$a_1 > a_2 > 0$$

$$b_1 > 0 > b_2$$

$$c_2 > c_1 > 0$$

$$d_2 > 0 > d_1 \quad .$$

Definition 2.2.1. *A worker w is* **acceptable** *to firm f if f prefers w to remaining unmatched.*

The second definition modifies the standard notion of an acceptable worker to account for the preferences of the firms at a given salary level.

Definition 2.2.2. *Given any salary, x_f, chosen by firm f, a worker w is* **salary-acceptable** *to firm f if f's value for worker w is greater than x_f.*

The values corresponding to each type are common knowledge, however, each firm knows only its own type. The types are drawn independently according to a common prior distribution π over $\mathcal{P}^f = \{a, b, c, d\}$. Given the two disjoint sets of agents, define a matching as follows.

Definition 2.2.3. *A matching is a function $\mu : F \cup W \to F \cup W \cup \emptyset$ such that*

(1) $\mu(f) \in W \cup \emptyset$ for all $f \in F$,

(2) $\mu(w) \in F \cup \emptyset$ for all $w \in W$, and

(3) $\mu(\mu(i)) = i$ for all $i \in F \cup W$ with $\mu(i) \neq \emptyset$.

Let $\mathcal{M}$ denote the set of all matchings.

For any firm f with type $k = (k_1, k_2)$, the utility derived from a matching $\mu \in \mathcal{M}$ is given by

$$u_k^f(\mu) = \begin{cases} k_1 & \text{if } \mu(f) = w_1 \\ k_2 & \text{if } \mu(f) = w_2 \\ 0 & \text{if } \mu(f) = \emptyset \end{cases} .$$

Before the matching is determined, the firms each choose a salary. Then the matching outcome is determined as follows.

Step 1. Each firm makes an offer to (at most) one worker;

Step 2. Each worker tentatively accepts at most one offer, and rejects all others; $\vdots$

Step k. Any firm whose most recent offer was rejected may make the same salary offer to a worker who has not already rejected them;

Step k+1. Each worker tentatively accepts at most one offer out of the one (if any) it tentatively holds, and the new offers received at *Step k*, and rejects all others.

The procedure terminates when no new offers are made, and then all tentative matches are confirmed.

In principle, both the firms and workers could adopt a large number of different strategies, some of which may be incredibly complex. However, the following two remarks, which are standard results from the matching theory literature, allow us to ignore any strategic behavior at the matching stage.

Remark 2.2.4. *For any set of chosen salaries, each firm has a dominant strategy to make offers in order of preference to salary-acceptable workers only.*

Remark 2.2.5. *For any profile of firm preferences and any set of chosen salaries, each worker has a dominant strategy to reject all but the highest salary offered to them.*

Since there are no incentives for strategic sequencing of offers by the firms, or for strategic rejection by the workers, the rest of the paper focuses on the behavior of the firms when they decide upon a salary. In fact, it is useful to describe the outcomes from the matching process by a direct revelation outcome function g. Let $g : \mathcal{P} \times \mathbb{R}^2_+ \to \mathcal{M}$ be an outcome function that maps the preferences (types) of the two firms and the salaries chosen by the firms into the set of matchings. Firms have a dominant strategy to announce their true preferences over salary-acceptable workers and the workers simply reject all but the highest offer made to them.

2.2.1 Pure Strategy Equilibria

Consider the game $\Gamma = \left(F, W, \mathcal{P}, \mathbb{R}_+, \pi, \mathbf{g}, \{u^f_k\}_{f,k}\right)$, which consists of the sets of firms F, and workers W, the firm type space $\mathcal{P}$, the space of possible salaries $\mathbb{R}_+$, and the type distribution π. The outcome function $\mathbf{g}$ represents the matching process described by Steps $1, 2, ...k+1$, and $\{u^f_k\}_{f,k}$ are the utility functions for each firm and each firm type over the set of matchings.

A pure strategy for a firm f is a function $\mathbf{s}_f : \mathcal{P}^f \to \mathbb{R}_+$ which selects a salary for each possible firm type. Given a strategy $\mathbf{s}_{-f}$ for the other firm, firm f's expected payoff from announcing a salary x_f when its type is k is given by

$$\mathbb{E}U^f_k\left(x_f, \mathbf{s}_{-f},\right) = \sum_{p \in \mathcal{P}^{-f}} \pi(p) \cdot u^f_k\left[g\left(k, p, x_f, s_{-f}(p)\right)\right].$$

The following arguments establish that there is no pure strategy Bayesian Nash equilibrium to the game, Γ.

Consider any arbitrary pair of strategies $(\mathbf{s}_1, \mathbf{s}_2)$ and suppose firm 1's type is a. Notice that, if firm 2's type is either of c or d, then regardless of $s_1(a)$, firm 1 is matched with worker w_1. However, if firm 2's type is a or b, then the outcome depends on the salaries announced by the firms. If firm 2 is playing $\mathbf{s}_2$, then firm 1's best response is to announce $s_1(a) = max\{s_2(a), s_2(b)\} + \varepsilon$ as long as $s_1(a) \leq a_1 - a_2$. If $max\{s_2(a), s_2(b)\} \geq a_1 - a_2$, then firm 1's best response is to announce $s_1(a) = 0$. However, given the choice of firm 1, $s_1(a) = max\{s_2(a), s_2(b)\} + \varepsilon$, firm 2's best response, if it is type a, is to offer $s_2(a) = s_1(a) + \varepsilon$, up to $s_2(a) \leq a_1 - a_2$. The same type of incremental best responses exist for type b firms, and by symmetry, also for types c and d. Thus, the problem with pure strategies is that firms who have a common most preferred worker will continue to outbid each other until the marginal benefit of 'winning' the worker is equal to the marginal benefit of not winning. However, once that point is reached, the best response is to announce a salary of 0, and begin the bidding-up process all over again.

2.2.2 Distributional (Mixed) Strategy Equilibria

Formally, a mixed (or distributional) strategy for firm i is a function $\sigma_i : \mathcal{P} \to \Delta(\mathbb{R}_+)$ which announces, for each preference type, a distribution over salaries in $\mathbb{R}_+$. For simplicity, refer to the symmetric equilibrium (σ^*, σ^*) by the equilibrium strategy $\sigma^* = (G_a^*, G_b^*, G_c^*, G_d^*)$ where G_k^* is the cumulative distribution announced by a firm whose type is k. I assume that strategies are continuous distributions with interval support.[5] Before turning to the results, it is useful to work through a simple example for the two-firm, two-worker model.

Example 2.2.6. *Suppose $a = (2, 1)$, $b = (2, -1)$, $c = (1, 2)$, and $d = (-1, 2)$, while $\pi(a) = \frac{1}{2}$, $\pi(b) = \frac{1}{8}$, $\pi(c) = \frac{1}{4}$, and $\pi(d) = \frac{1}{8}$. Notice that the marginal benefit to getting worker w_1 is higher for type b than type a, and the marginal benefit to getting worker w_2 is higher for type d than type c. Given these parameters, a natural conjecture is that type b firms will make higher offers than type a firms, and type d firms will make higher offers than type c firms. Furthermore, given the distribution of types π, worker w_1 is ex ante more popular (or believed to be more popular) than w_2. As such, one might expect to see higher salaries on average being offered to w_1.*

For this example, there exists an equilibrium, which is described as follows:

$$\begin{aligned}
G_a^*(x) &= 2x \quad \textit{on the support } \left[0, \frac{1}{2}\right] \\
G_b^*(x) &= \frac{7x - 3.5}{2 - x} \quad \textit{on the support } \left[\frac{1}{2}, \frac{11}{16}\right] \\
G_c^*(x) &= 4x \quad \textit{on the support } \left[0, \frac{1}{4}\right] \\
G_d^*(x) &= \frac{7x - 1.75}{2 - x} \quad \textit{on the support } \left[\frac{1}{4}, \frac{15}{32}\right].
\end{aligned}$$

This equilibrium exhibits several interesting features. First, there is no overlap between the

[5] There may be other types of symmetric equilibria, with noninterval support, or discontinuous strategies. In addition, there may be asymmetric equilibria.

equilibrium supports of types with a common most preferred worker. Since the marginal value of getting worker w_1 is less for type a than for type b, firms of type b **always** announce higher salaries than firms of type a. In other words, firms of type a are resigned to getting their second favorite worker (w_2) when the other firm is type b. Instead, a type a firm focuses just on competing against another type a firm.

On the other hand, a type b firm offers enough to ensure that it outbids any type a firm, then focuses on competing against the chance that the other firm is a type b. This type of'separation result between types a and b is also exhibited by types c and d, and as will be shown in Section 2.3, is a characteristic of any equilibrium in continuous distributional strategies with interval support.

Second, equilibrium salaries are higher on average for firms of type a than type c and for type b than type d, even though they have comparable values for their respective preferences. This reflects the relative popularity of worker w_1 over worker w_2. This notion of popularity is manifested in the differences in the probabilities of facing another firm with the same most preferred worker. For types a and b, the probability of facing another type a or b is $\frac{5}{8}$, while for types c and d, the probability of facing another type c or d is only $\frac{3}{8}$. As a result, the average salaries offered in equilibrium are higher for type a than type c, and higher for type b than type d. Section 2.3 confirms that this feature is a general result that applies to all equilibria of the game.

2.3 Equilibria in the Two-Firm, Two-Worker Model

Consider the general two-firm, two-worker model. It is relatively straightforward to show that the strategies of types that share a common most preferred worker affect each other. On the other hand, salaries do not affect the matching output when the realized types do not have a common most preferred worker. Thus, pairs of types with common most preferred workers can be considered in isolation from one another. Without loss of generality, consider types a and b. The following two lemmas allow us to characterize the supports for the equilibrium strategies.

Lemma 2.3.1. *Between types with a common most preferred worker, the lowest salary offered in* *equilibrium must be 0.*

Proof 1. *Let* $[\underline{x}_a, \overline{x}_a]$ *and* $[\underline{x}_b, \overline{x}_b]$ *be the equilibrium supports for types* a *and* b*, respectively. Suppose by means of contradiction that neither* $\underline{x}_a$ *nor* $\underline{x}_b$ *is equal to 0. Consider* $0 < \underline{x}_a \leq \underline{x}_b$*. Type* a*'s expected payoff from* $x = \underline{x}_a$ *is*

$$\mathbb{E}U_a(\underline{x}_a) = [\pi(a) + \pi(b)]a_2 + [\pi(c) + \pi(d)]a_1 - \underline{x}_a$$

and for any $x \in [0, \underline{x}_a)$*, type* a*'s expected payoff is*

$$\begin{aligned} \mathbb{E}U_a(x) &= [\pi(a) + \pi(b)]a_2 + [\pi(c) + \pi(d)]a_1 - x \\ &< \mathbb{E}U_a(\underline{x}_a). \end{aligned}$$

This means that $[\underline{x}_a, \overline{x}_a]$ *cannot be an equilibrium support unless* $\underline{x}_a = 0$ *or* $0 \leq \underline{x}_b < \underline{x}_a$*.*

If $0 < \underline{x}_b \leq \underline{x}_a$*, type* b*'s expected payoff from* $x = \underline{x}_b$ *is*

$$\mathbb{E}U_b(\underline{x}_b) = (b_1 - \underline{x}_b)[\pi(c) + \pi(d)].$$

That is, at the lower bound of type b*'s equilibrium support, a firm of type* b *does not get matched to a worker unless the other firm is type* c *or type* d*. But in those cases, the salary does not affect the outcome, so that choosing a salary of* $\underline{x}_b > 0$ *is strictly dominated by* $x = 0$*. Thus,* $[\underline{x}_b, \overline{x}_b]$ *cannot be an equilibrium support unless* $\underline{x}_b = 0$ *or* $0 \leq \underline{x}_a < \underline{x}_b$*. Therefore, in equilibrium, we must have either* $\underline{x}_a = 0$ *or* $\underline{x}_b = 0$*.*

Lemma 2.3.2. *In equilibrium, there are no gaps between the equilibrium supports for types with a common most preferred worker.*

Proof 2. *Suppose* $\overline{x}_a < \underline{x}_b$*. Then* $\forall x \in (\overline{x}_a, \underline{x}_b)$*, type* b*'s expected payoff is*

$$\begin{aligned}\mathbb{E}U_b(x) &= (b_1 - x)(1 - \pi(b)) \\ &> (b_1 - \underline{x}_b)(1 - \pi(b)) = \mathbb{E}U_b(\underline{x}_b),\end{aligned}$$

contradicting the inclusion of $\underline{x}_b$ *in the equilibrium support for type* b*. The proof is similar for the case when* $\overline{x}_b < \underline{x}_a$*. Since the supports are intervals by assumption, there are no other cases to be considered.*

These two lemmas imply that equilibria must be consistent with one of four cases. In each case, type a mixes over $[\underline{x}_a, \overline{x}_a]$, and type b mixes over $[\underline{x}_b, \overline{x}_b]$, where

Case 1. $0 = \underline{x}_a < \underline{x}_b \leq \overline{x}_a < \overline{x}_b$

Case 2. $0 = \underline{x}_b < \underline{x}_a \leq \overline{x}_b < \overline{x}_a$

Case 3. $[\underline{x}_a, \overline{x}_a] \subset [\underline{x}_b, \overline{x}_b]$, and $\underline{x}_b = 0$

Case 4. $[\underline{x}_b, \overline{x}_b] \subset [\underline{x}_a, \overline{x}_a]$, and $\underline{x}_a = 0$.

Proposition 2.3.3 generalizes and formalizes the separation result illustrated in Example 2.2.6, by showing that there cannot be equilibria of the form described by *Case 3* or *Case 4*.

Proposition 2.3.3. *Equilibrium supports do not overlap for types with a common most preferred worker. In particular then, all equilibria must be of the form in Case 1 with* $\underline{x}_b = \overline{x}_a$ *or Case 2 with* $\underline{x}_a = \overline{x}_b$.

Proof 3. *See Appendix.*

The proof for Proposition 2.3.3 is based on demonstrating that indifference cannot be satisfied simultaneously for both types on an interval with nonempty interior. As a result, the equilibrium supports in *Case 1* and *Case 2* must meet at their boundaries. For *Case 3* and *Case 4*, the same argument implies that a support which is a subset of the other must be a single point. Since best responses in pure strategies have already been ruled out, there must not exist an equilibrium in which one support is nested in the other.

The second result characterizes all symmetric Bayesian Nash equilibria in which strategies are continuous with interval support. Moreover, it provides the set of conditions that determine, for each pair of types with a common most preferred worker, whether their equilibrium supports are consistent with *Case 1* or *Case 2*. The condition depends on the relative marginal benefits of getting the types' common most preferred worker, and on the probability that a firm is the type that also finds the other worker acceptable.

Proposition 2.3.4. *(i) If $b_1 > \pi(a)(a_1 - a_2)$, then in all equilibria,*

$$\begin{aligned} G_a^*(x) &= \frac{x}{\pi(a)(a_1 - a_2)} \\ &\quad \textit{on the support } \Big[0,\, \pi(a)(a_1 - a_2)\Big] \\ G_b^*(x) &= \frac{1 - \pi(b)}{\pi(b)(b_1 - x)}\left[x - \pi(a)(a_1 - a_2)\right] \\ &\quad \textit{on the support } \Big[\pi(a)(a_1 - a_2),\, \pi(b)b_1 + (1 - \pi(b))\pi(a)(a_1 - a_2)\Big], \end{aligned}$$

regardless of G_c^*, G_d^*. The analogous result holds for types c and d if $d_2 > \pi(c)(c_2 - c_1)$.

(ii) If $b_1 < \pi(a)(a_1 - a_2)$, then in all equilibria,

$$\begin{aligned} G_b^*(x) &= \frac{x(\pi(c) + \pi(d))}{\pi(b)(b_1 - x)} \\ &\quad \textit{on the support } \left[0,\, \frac{\pi(b)b_1}{1 - \pi(a)}\right] \\ G_a^*(x) &= \frac{(1 - \pi(a))x - \pi(b)b_1}{\pi(a)(1 - \pi(a))(a_1 - a_2)} \\ &\quad \textit{on the support } \left[\frac{\pi(b)b_1}{1 - \pi(a)},\, \frac{\pi(b)b_1}{1 - \pi(a)} + \pi(a)(a_1 - a_2)\right]. \end{aligned}$$

The analogous result holds for types c and d if $d_2 < \pi(c)(c_2 - c_1)$.

Proof 4. *See Appendix.*

Part of the condition in Proposition 2.3.4 has a simple intuition. If type b gets a higher value from worker w_1 than the marginal value for type a from getting w_1 instead of w_2, then type b will be willing to pay more than type a for w_1. The role of $\pi(a)$ in the condition is less obvious. Keeping the values fixed, if $\pi(a)$ is relatively low, a type a firm does not need to mix over a large interval to compete against its own type. As a result, if type a firms offer salaries above those offered by type b, there may be an incentive for type b firms to offer salaries higher than the type a firms in order to 'steal' worker w_1 in the event that the other firm is type a. Any such deviation by type b firms would give type a firms an incentive to lower the support of their distributional strategies to the lower bound of 0.

Proposition 2.3.4 also leads to two corollaries. First, all things being equal, the more likely a firm is to face another firm of the same type, the stronger the competitive pressure and the higher the average equilibrium salary offered by that type. Similarly, the more likely a firm is to face another firm with the same most preferred worker, the stronger the competition and the higher the average equilibrium salary offered by the two relevant types.

Corollary 2.3.5. *The higher the probability a firm type has to compete against its own type, the higher (on average) the equilibrium salary offered by that firm type.*

Corollary 2.3.6. *For any firm, the higher the probability that the other firm has the same most preferred worker, the higher the equilibrium salary (on average) offered by the firm.*

2.4 Competition in Large Markets

This section examines equilibrium behavior in large markets. For tractability, I replicate the two-firm, two-worker market to obtain a market with $2n$ firms and $2n$ workers, consisting of n identical class w_1 workers and n identical class w_2 workers. I characterize the equilibria for the case in which there are a continuum of firms and a continuum of workers, then prove existence, for finite replicated markets, of a Bayesian Nash equilibrium in continuous distributional strategies with interval supports. Moreover, both Proposition 2.3.3 and Proposition 2.3.4 generalize to replicated markets.

2.4.1 Market Replication

Using replicated markets avoids the problem of having an exponentially growing type space. Replication provides a convenient way to conduct a tractable analysis of competitive behavior in large markets. Furthermore, it removes that chance of realizing an uninteresting market with sparse competition, in which every firm desires a different type of worker.

The baseline market is the two-firm, two-worker market, with $F^1 = \{f_1, f_2\}$ and $W = \{w_1, w_2\}$. In an n-replicated market, there are $2n$ firms, $F^n = \{f_1, ..., f_{2n}\}$, along with n copies of w_1, $W_1 = \{w_1^1, w_1^2, ..., w_1^n\}$, and n copies of w_2, $W_2 = \{w_2^1, ..., w_2^n\}$. Since w_1^j and w_1^k are identical copies of one another, we assume that all firms are indifferent between any two workers in W_1. Likewise, all firms are indifferent between any two workers in W_2. As a result, firms' preferences (and from these, their types) are defined as strict orderings over the set $\{W_1, W_2, \emptyset\}$.

As in Section 2.2, firm types that prefer being unmatched over every worker are ignored.[6] This leaves four possible firm types that are essentially the same as the types in the two-firm, two-worker model, except that the preferences are over classes of workers W_1 and W_2.

$$
\begin{array}{ll}
P_a : \; W_1 \, W_2 \, \emptyset & \qquad P_b : \; W_1 \, \emptyset \, W_2 \\
P_c : \; W_2 \, W_1 \, \emptyset & \qquad P_d : \; W_2 \, \emptyset \, W_1 .
\end{array}
$$

Assume that each preference ranking is represented by a pair of values - one for each worker class, W_1 and W_2 - while the value of remaining unmatched is normalized to 0. So, for each type $k \in \{a, b, c, d\}$, $k = (k_1, k_2)$, where k_i is the value of each worker w in the class W_i. For the values to represent the

[6] We may just as well assume that they don't enter the market in the first place.

corresponding preference rankings, they must satisfy

$$a_1 > a_2 > 0$$

$$b_1 > 0 > b_2$$

$$c_2 > c_1 > 0$$

$$d_2 > 0 > d_1 \quad .$$

Each firm knows only its own type, and the types are drawn independently according to the common prior distribution π over $\{a, b, c, d\}$. That is, $\pi(k)$ is the probability that a given firm is a type k firm, or equivalently, has preferences P_k.

2.4.2 Equilibria in the Continuum Case

Before analyzing the equilibria for a finite replicated market, consider the equilibrium behavior in the limit, when there is a continuum of firms, and continuum of workers. Moreover, suppose that the measure of workers in each class W_1 and W_2 is half the total measure of W. In this environment, since there are infinitely many firms, the aggregate uncertainty about the realized firm types disappears from the market. That is, $\pi(k)$ is the actual proportion, or the measure of type k firms in the market. This is a convenient feature because it makes the equilibrium strategies relatively straightforward functions of the distribution π.

As for the two-firm, two-worker case, the equilibrium strategy for a given type k does not depend on the strategies of the two types k', k'' that have a different most preferred worker class than type k. Thus, as in Section 2.2, when deriving equilibrium strategies, types a and b can be treated independently from types c and d. Thus, without loss of generality, consider types a and b. The analysis is symmetric for types c and d. Proposition 2.4.1 characterizes the equilibria for the limit case. It is broken into two cases based on the relative marginal values of worker class W_1 compared with worker class W_2, for types a and b.

Proposition 2.4.1. *The following two cases characterize the equilibria when there are a continuum of firms and a continuum of workers, with two equally large worker classes.*

(1) $b_1 \geq a_1 - a_2$

(i) If $\pi(a) + \pi(b) \leq \frac{1}{2}$*, then* $x_a^* = 0$ *and* $x_b^* = 0$.

(ii) If $\pi(a) > \frac{1}{2}$*, then* $x_b^* = a_1 - a_2$ *and*

$$x_a^* = \begin{cases} 0 & \textit{with probability } p_a(0) = \frac{2(\pi(a)+\pi(b))-1}{2\pi(a)} \\ a_1 - a_2 & \textit{with probability } 1 - p_a(0) \end{cases}.$$

(iii) If $\pi(b) > \frac{1}{2}$*, then* $x_a^* = 0$ *and* $x_b^* = b_1$.

(iv) If $\pi(a) \leq \frac{1}{2}$*,* $\pi(b) \leq \frac{1}{2}$*, but* $\pi(a) + \pi(b) > \frac{1}{2}$*, then* $x_b^* = a_1 - a_2$ *and*

$$x_a^* = \begin{cases} 0 & \textit{with probability } p_a(0) = \frac{2(\pi(a)+\pi(b))-1}{2\pi(a)} \\ a_1 - a_2 & \textit{with probability } 1 - p_a(0) \end{cases}.$$

(2) $b_1 < a_1 - a_2$

(i) If $\pi(a) + \pi(b) \leq \frac{1}{2}$*, then* $x_a^* = 0$ *and* $x_b^* = 0$.

(ii) If $\pi(a) > \frac{1}{2}$*, then* $x_b^* \in [0, b_1]$ *and*

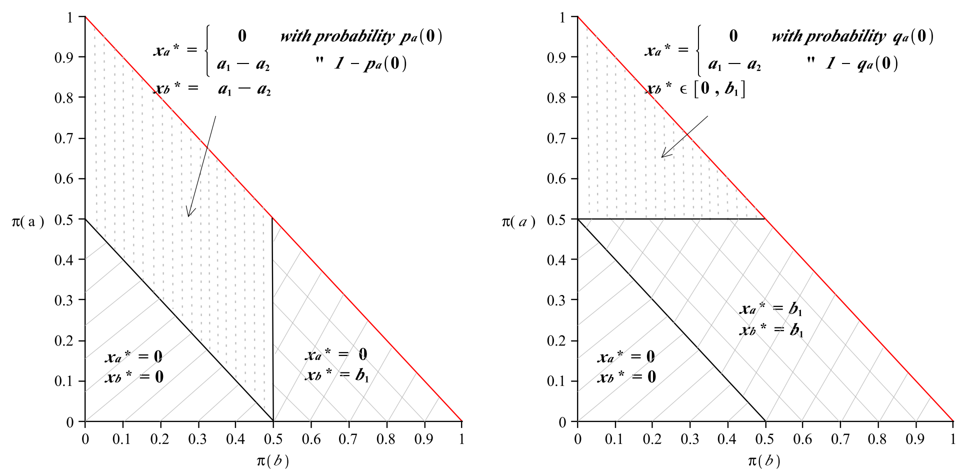
$$x_a^* = \begin{cases} 0 & \textit{with probability } q_a(0) = \frac{2\pi(a)-1}{2\pi(a)} \\ a_1 - a_2 & \textit{with probability } 1 - q_a(0) \end{cases}.$$

(iii) If $\pi(b) > \frac{1}{2}$*, then* $x_a^* = b_1$ *and* $x_b^* = b_1$.

(iv) If $\pi(a) \leq \frac{1}{2}$*,* $\pi(b) \leq \frac{1}{2}$*, but* $\pi(a) + \pi(b) > \frac{1}{2}$*, then* $x_a^* = b_1$ *and* $x_b^* = b_1$.

The proof, which is quite straightforward, is omitted. Instead, Figures 2.1a and 2.1b provide graphical illustrations of the two cases in Proposition 2.4.1. Each figure plots $\pi(a)$ against $\pi(b)$ and divides the space of probability pairs $(\pi(b), \pi(a))$ into segments for each subcase of the equilibrium characterization. In both Figure 2.1a and Figure 2.1b, the bottom-left triangle corresponds to the case in which there is an excess supply of class W_1 workers, and therefore no competition between types a and b. Thus, $x_a^* = x_b^* = 0$ for both cases when $\pi(a) + \pi(b) \leq \frac{1}{2}$.

Figure 2.1a merges the subcase in which $\pi(a) > \frac{1}{2}$ with the subcase in which $\pi(a) \leq \frac{1}{2}$ and $\pi(b) \leq \frac{1}{2}$, but $\pi(a) + \pi(b) > \frac{1}{2}$, since in each, type a firms mix between 0 and $a_1 - a_2$ with probability $p_a(0) = \frac{2[\pi(a)+\pi(b)]-1}{2\pi(a)}$, while type b firms choose $a_1 - a_2$. Finally, in the case when $\pi(b) > \frac{1}{2}$, type b firms compete with each other and push the salary up to their marginal value from a class W_1 worker, while type a firms know that they will not be matched with a class W_1 worker and so choose a salary of 0.



(a) Case (1): $b_1 \geq a_1 - a_2$

(b) Case (2): $b_1 < a_1 - a_2$

Figure 2.1: Continuum equilibria for firm types a and b in Cases (1) and (2) from Proposition 2.4.1

In Figure 2.1b, we can likewise merge the subcase in which $\pi(b) > \frac{1}{2}$ with the subcase in which $\pi(a) \leq \frac{1}{2}$ and $\pi(b) \leq \frac{1}{2}$, but $\pi(a) + \pi(b) > \frac{1}{2}$, since in each subcase, both type a firms and type b firms choose a salary of b_1. When $\pi(a) > \frac{1}{2}$, type a firms mix between 0 and $a_1 - a_2$ with probability $q_a(0) = \frac{2\pi(a)-1}{2\pi(a)}$, while type b firms choose a salary in the interval $[0, b_1]$. This is because type a firms drive the salary for a class W_1 worker up to $a_1 - a_2 > b_1$, so that type b firms are never matched with anyone. Since some of the type a firms will miss out on a class W_1 worker, they mix between the salary $a_1 - a_2$ and 0.

2.4.3 Finite Replicated Markets

As in both the two-firm, two-worker and the continuum cases, the equilibrium depends on the parameters of the model. We can break up the proof of existence into several cases. The proof is by construction. We consider firm types a and b, although things are symmetric for types c and d.

Proposition 2.4.2. *Given any finite replicated market with* $2n$ *firms,* n *workers in class* W_1 *and* n *workers in class* W_2*, there exists an equilibrium* $(G_a^*(\cdot), G_b^*(\cdot), G_c^*(\cdot), G_d^*(\cdot))$ *such that* $G_k^*(\cdot)$ *is a continuous distribution with interval support in the salary space, for all* $k = a, b, c, d$*. The equilibrium supports for types* a *and* b *satisfy*

$$\begin{aligned} 0 = \underline{x}_a < \overline{x}_a &= \underline{x}_b < \overline{x}_b \\ \text{or} \quad 0 = \underline{x}_b < \overline{x}_b &= \underline{x}_a < \overline{x}_a. \end{aligned}$$

The analogous result holds for the equilibrium supports of types c *and* d*.*

Proof 5. *See Appendix.*

2.5 Convergence of Finite Market Equilibria

This section shows numerically that the replicated market equilibrium strategies converge to the corresponding continuum equilibrium as the number of replications goes to infinity. The convergence is illustrated by simulating replicated markets for the market presented in Example 2.2.6. Recall that $a = (2, 1)$, $b = (2, -1)$, $c = (1, 2)$, and $d = (-1, 2)$, while $\pi(a) = \frac{1}{2}$, $\pi(b) = \frac{1}{8}$, $\pi(c) = \frac{1}{4}$, and

$\pi(d) = \frac{1}{8}$. The corresponding continuum equilibrium is as follows,

$$x_a^* = \begin{cases} 0 & \text{with probability } \frac{1}{4} \\ 1 & \text{with probability } \frac{3}{4} \end{cases} \tag{2.1}$$

$$x_b^* = 1 \tag{2.2}$$

$$x_c^* = 0 \tag{2.3}$$

$$x_d^* = 0. \tag{2.4}$$

The equilibrium distribution for a type a firm in an n-replicated market satisfies

$$\underline{x}_a = 0$$

$$\overline{x}_a = (a_1 - a_2)\left[\sum_{j=0}^{n-1} \sum_{k=n+1-j}^{2n-1-j} \frac{(2n-1)!\pi(b)^j \pi(a)^k [1 - \pi(b) - \pi(a)]^{2n-1-j-k}}{j!k!(2n-1-j-k)!}\right]$$

and for all $x \in [\underline{x}_a, \overline{x}_a]$,

$$x = (a_1 - a_2) \sum_{j=0}^{n-1} \sum_{k=n+1-j}^{2n-1-j} \left[\frac{(2n-1)!\pi(b)^j \pi(a)^k [1 - \pi(a) - \pi(b)]^{2n-1-j-k}}{j!k!(2n-1-j-k)!} \times \sum_{t=0}^{n-j-1} \binom{k}{t} G_a^*(x)^{k-t} [1 - G_a^*(x)]^t \right].$$

The last equation can be solved, given any value of $G_a^*(x)$, for the corresponding value of x. The pairs $(x, G_a^*(x))$ that satisfy the indifference equations for a given value of n can be used to trace out the equilibrium distribution for a type a firm in a market that has been replicated n times. Figure 2.2a plots these pairs $(x, G_a^*(x))$ with x on the horizontal axis and $G_a^*(x)$ on the vertical axis for several different values of n.

For smaller sized markets, $n \in \{2, 3, 6, 10, 20\}$, increasing the market size shifts more density to higher salaries and expands the equilibrium support. However, after n grows large enough, the equilibrium support approaches its upper bound of 1 (equal to $a_1 - a_2$). Then for any larger replicated markets, type a firms shift greater weight towards salaries very close to the upper bound. However, in order to maintain indifference over the support, they also assign greater probability to very low salaries (close to 0), which leads to a CDF that approaches the continuum equilibrium as n approaches infinity.

The same procedure can be run for Type b firms, and also for Types c and d. For type b firms, the corresponding equilibrium support has a lower bound equal to the upper bound of the support for type a firms, by the separation result. The size of the support for a type b firm's equilibrium strategy is decreasing with the size of the market, and since the lower bound for sufficiently large n is equal to 1, the equilibrium strategies converge towards the continuum equilibrium strategy, which places the entire mass on a salary equal to 1. Figure 2.3 shows the simulated calculations for equilibrium strategies of a type b firm in different sized markets. Of particular interest is the observation that by $n = 6$, there is already almost no competitive pressure for a type b firm to compete against another type b firm. Even with so few replications, a type b firm realizes that the chances of there being more than 5 other type b firms (among the other 11 firms) is very low. As a result, type b firms choose a salary just high enough to ensure that they will be ranked higher by the workers than any type a firms.

For types c and d, as the market is replicated, firms perceive that the chances of excess demand for W_2 class workers are very low. It is very easy to show that the equilibrium distributions shift towards the corresponding continuum equilibrium strategies, which place the entire mass on a salary of 0. In this respect, as the market is replicated, competitive pressures are enhanced only for the most popular worker class. However, even then, there is no pressure to compete applied to firms of the type that has a higher relative marginal value for the popular workers.

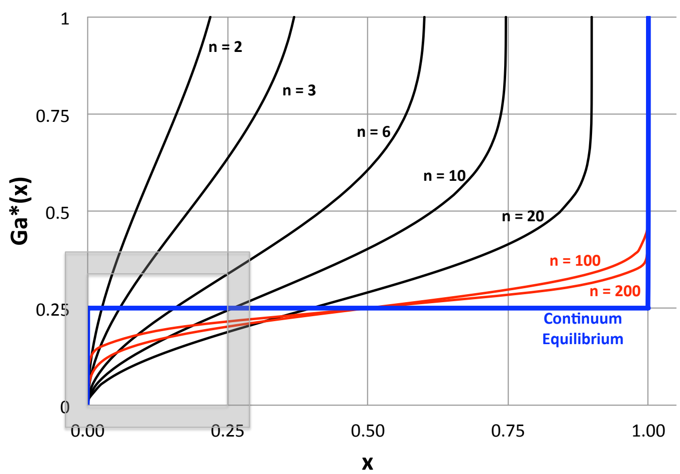
2.6 Conclusion

This paper presents three main results. First, the equilibrium characterization for the two-firm, two-worker model shows that strategy supports for types with a common most preferred worker are adjacent, such that the type with a higher relative marginal value for the preferred worker pays a premium to ensure it is preferred over any firm of the other type. Second, in the limit, when there are a continuum of firms and a continuum of workers, there is no aggregate uncertainty about the realization of firm types. As a result, competition is confined to the class of workers that are more popular, while the salaries of other workers fall to zero. Third, when the two-firm, two-worker model is replicated to form larger markets, there exist distributional Bayesian Nash equilibrium strategies that exhibit the same separation result in the equilibrium supports for types with a common most preferred worker class. Numerical results suggest that, as the number of replications increases, the equilibrium strategies in the finite replicated market approach the corresponding continuum equilibrium strategies. As a result, when markets become larger, aggregate uncertainty about the actual types of other firms dissipates and reduces the level of competitive pressure on salaries.

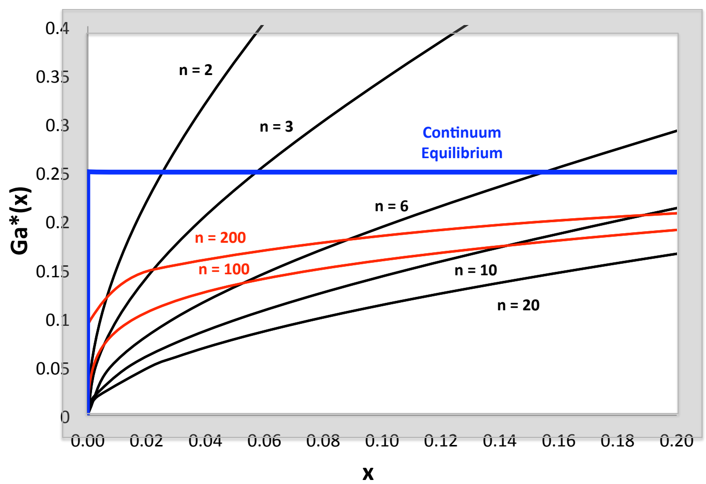
A natural extension of this paper is to consider a general n-firm, n-worker model. In this paper, large markets are generated by replicating the two-firm, two-worker model, which controls the size of the type space and keeps the analysis tractable. A limitation of this approach is that all workers in the same class are treated as identical from the perspective of the firms. Future work might focus on relaxing this assumption, while maintaining tractability. An alternative approach may be to relax the restriction that firms with the same preference ordering must have the same valuation or utility representation for that ordering. For instance, suppose each firm's type is a pair of values $\theta = (x, y)$, each drawn independently from some interval $[\underline{\theta}, \overline{\theta}]$ according to a given distribution.

Both the two-firm, two-worker model, or a small $n \times n$ model with just two worker classes, are simple and interesting candidates for experimental work. Although in theory firms have a dominant strategy to make offers in order of preference, experimental work will reveal the actual sequence of offers chosen by the subject firms. Experimental data might provide some indication of the strategies firms actually play when their types are private information.

Finally, it may prove useful to examine how the theoretical results in this paper fit with empirical observations in different labor markets. In many labor markets, salaries do not accurately reflect differences in productivity. Instead, salaries tend to exhibit less variation than worker productivities. Although the model in this paper does not include explicit levels of productivity, a modified version may provide some explanation for this empirical trend in workers' salaries.



(a) Type a firms' equilibrium strategies for $n \in \{2, 3, 6, 10, 20, 100, 200\}$



(b) Magnified view of the salary space for $x \in [0, 0.2]$

Figure 2.2: Replicated market equilibria for type a. In panel (a), each curve shows the equilibrium CDF (cumulative distribution function) for a type a firm, for a given market size. Panel (b) shows a magnified view of the CDFs at low salaries to illustrate how the density decreases with n in relatively small markets, then increases with n for sufficiently large markets.

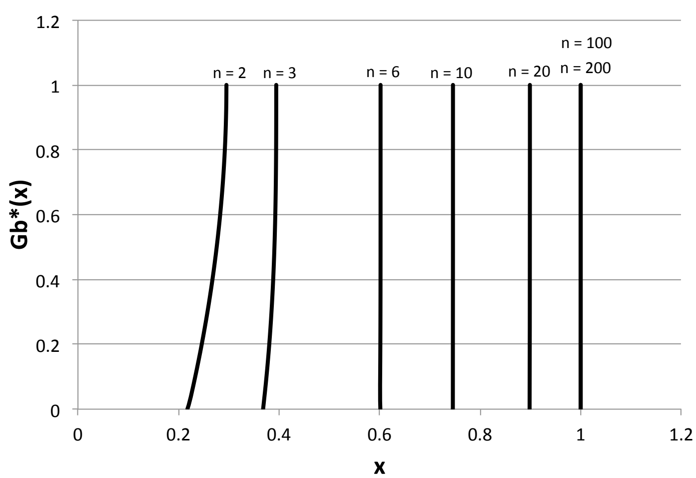


Figure 2.3: Replicated market equilibria for type b

Chapter 3

Information and Cooperation in a Network Public Goods Experiment

3.1 Introduction

There are many different settings in which public goods are provided using the voluntary contributions mechanism. Examples include fundraising by local district school boards, charitable contributions, or political campaign funds. There is a vast body of theoretical and experimental literature focused on explaining why individuals contribute in these environments, despite incentives to free-ride. Much of this literature studies the finitely repeated setting in which players are given endowments of a private good, which they may contribute towards the production of a linear public good. In some environments, individuals are active members in several different communities, and their decisions or actions affect outcomes for multiple neighborhoods. These types of environments have motivated recent research into public goods on networks, beginning with theoretical work by Bramoullé and Kranton (2007).

Experimental evidence suggests a robust pattern of contributions in standard repeated public goods games. Average contributions in the first period are typically around half of the endowment, then decline steadily over time, although they do not generally converge to zero. This pattern is well documented in the literature, beginning with Isaac, Walker, and Thomas (1984) and Isaac, McCue, and Plott (1985). A second key finding in the experimental literature is that many participants are conditional cooperators whose contributions to the public good are positively correlated with the contributions (or their beliefs about the contributions) made by others. Chaudhuri (2011) provides an excellent recent summary of the literature that aims to identify and explain conditional cooperation.[1]

The study of conditional cooperation raises two important questions, which provide the motivation for this paper. First, *how does conditional cooperation depend on the nature of the information players observe about others between rounds of the repeated game?* It seems natural to expect that this feature of the environment will affect both *how*, and *on what* the players condition their contribution decisions. Moreover, this type of information can vary significantly in public good settings. For instance, school district boards tend to publish aggregated statistics about contributions, broken down by class, neighborhood, or level of contribution. In other cases, participants may be able to observe the individual contributions made by others.

Second, *how does conditional cooperation depend on the structure of interaction between players?* In particular, if players are connected in a network, how do their positions affect contributions? Bramoullé and Kranton (2007) cite several classic examples of goods that are nonexcludable along social or geographical links, including innovation, or pollution abatement. In these types of environments, information about participants with whom the subject is not connected may nevertheless be relevant to the way he chooses to behave. For instance, a conditional cooperator may find it relevant how much his neighbors' neighbors contributed toward the public good.

We investigate these two questions using a series of experiments, in which the subjects play a finitely repeated linear public goods game on a circle network. Each player's consumption of the public good depends only on his own contribution, and on those of his neighbors in the network. The information players are shown between rounds is controlled in two ways. The first exploits the network structure of the environment, by varying whether subjects observe information about the whole population or only about their neighborhood. We call this aspect of the design the *treatment group*. The second variation changes the type of information statistic shown to subjects about the relevant *treatment group*. Specifically, we vary the statistic between average contribution, and

[1] Some of the main contributions to the literature include Fischbacher, Gachter, and Fehr (2001); Fischbacher and Gachter (2010); Croson, Fatas, and Neugebauer (2005); Croson (2007); Frey and Meier (2004); Keser and van Winden (2000).

average payoff, for the given treatment group.[2]

We propose a behavioral model that extends the approach used by Arifovic and Ledyard (2009) to study contributions in a standard, complete network, repeated public goods game. They combine their Individual Evolutionary Learning (IEL) model with a relatively simple model of outcome-based other-regarding preferences (ORP). The model assumes that players have preferences which reflect altruism and envy, in the form of a concern for fairness with respect to one's self. We allow for heterogeneity across players through their types, which consist of a pair of parameters representing the player's degrees of altruism and envy. Each player assesses his utility on the basis of a particular *reference statistic*, with respect to a specific *reference group* of other players. For example, players may derive utility from an increase in the average payoff earned by their reference group. On the other hand, they may derive disutility if their own payoff is less than the average payoff in their reference group, due to envy. This model is most similar to the class of outcome-based ORP models, which includes the models developed by Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002).

Our approach to preferences is fairly standard among models of other-regarding behavior. However, there are several novel features that we incorporate in order to extend upon the standard approach. First, we allow explicitly for variation in a player's reference group. This is an integral component of other-regarding preferences that is typically underemphasized by assuming that the reference group is 'everyone else'.

Second, we consider two different possible criteria or statistics that players may use to make the social comparisons that inform their other-regarding preferences. Specifically, when reference groups are only a subset of the entire population, whether the players care about others' contributions or others' payoffs may lead to different behavior. For a fixed reference group, a player concerned with payoffs cares about the contributions made by the members of his reference group (his neighbors), but also the contributions made by his neighbors' neighbors, some of who may be outside his reference group. To this point, our approach is grounded in standard other-regarding preferences, albeit with

[2] In the circle network, these statistics are not informationally equivalent for the neighborhood treatment groups, since the average payoff conveys some additional information about the contributions made by a player's neighbors' neighbors.

a sharper focus on reference groups and statistics.

The less standard part of our approach is the claim that in the laboratory setting, we can control these two important features of the players' other-regarding preferences. Subjects may enter the lab with an underlying preference for altruism and fairness, but are less likely to show up with a predetermined reference group or reference statistic on which to condition their decisions. Since these reference-based features depend critically on the information made available to the subjects, we argue that we can manipulate players' reference groups and statistics, and perhaps then their decisions, by changing the information provided to them in the experiment.

An alternative, related approach to outcome-based other-regarding preferences argues that individuals reciprocate the behavior of others, cooperating with those who cooperate and not with those who don't. An important component of the theory of reciprocity is the idea that intentions, not just outcomes, matter for fairness, so that both kindness and unkindness are reciprocated. This approach is followed and developed by Rabin (1993); Levine (1998); Dufwenberg and Kirchsteiger (2004); Falk and Fischbacher (2006) and Ambrus and Pathak (2011). In recent years, this research has shown that incorporating intentions can help to explain features of conditionally cooperative behavior that are unaccounted for by outcome-based models. However, the experimental literature also suggests that intentions alone are insufficient to explain the scope of conditional cooperation. In this paper, we focus on a simpler, outcome-based model, which captures the essential features of conditional cooperation. Reciprocity models like those developed in Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) are equilibrium-based models that involve either a great deal of common knowledge about other players' preferences, or a well-defined set of ex-ante beliefs about every possible path of play in the finitely repeated game. Our simpler approach requires no such assumptions.

The primary innovation of the model is the assumed link between information treatment and behavior. Specifically, we assume that a player's reference group is precisely the *treatment group* about whom he is shown information, and that his reference statistic is precisely the *information statistic* that he is shown about the relevant treatment group between rounds. Thus, when the information treatment shows players the average payoff in their neighborhood, each player assesses outcomes with respect to the average payoff in his neighborhood. If they are instead shown the average contribution in the whole population, they assess utility with respect to the average contribution in the population. In this way, changing the information treatment changes the utility function employed by the players, by making certain outcomes and comparisons more salient.

The other important feature of the model is the learning hypothesis. We argue that players are boundedly rational agents who learn to play the game over time. Following Arifovic and Ledyard (2009), we use the model of Individual Evolutionary Learning (IEL) to represent the learning process. While there are several other widely studied learning models, such as Reinforcement Learning or Experience-Weighted Attraction, IEL is particularly well-suited to repeated games with large, continuous action spaces, like the ones considered in this paper.[3] The IEL model is reactive and individual-based, which allows us to avoid making restrictive assumptions about common knowledge or beliefs, and it has already been used successfully by Arifovic and Ledyard (2009) to explain a number of stylized facts from existing experimental data in linear public goods games. We use both the modified model of other-regarding preferences, which depend on the information treatment, and the IEL model, altered to incorporate the network structure of the game, to simulate the learning behavior of the subjects in the series of network public goods experiments.

While we focus on the learning hypothesis, there is another widely adopted approach to explaining the dynamics of average contributions in repeated games. The *strategic* approach contends that the decay in average contributions can be explained as strategic equilibrium behavior in which selfish players cooperate early in the game, in order to stimulate contributions from conditionally cooperative players. Although this approach is theoretically quite appealing, it requires a great deal of common knowledge about all players' preferences, beliefs, and information. While these assumptions may be appropriate in simple games like a repeated prisoner's dilemma, they are difficult to justify in a repeated public goods game with many periods and a continuous action space, let alone in such a game where outcomes and contributions are not always perfectly observed. Thus,

[3]For details of other learning models, see Roth and Erev (1995) for Reinforcement Learning, and Camerer and Ho (1999) for Experience-Weighted Attraction.

we adopt the learning hypothesis, in conjunction with a theory of other-regarding preferences, to model individual behavior in the finitely repeated linear public goods experiments.

The main contributions of the paper are as follows. First, we confirm a considerable amount of heterogeneity in contribution decisions across individuals. Some subjects contribute everything, some contribute nothing, while others contribute increasing or decreasing amounts between zero and full contribution. A significant proportion of subjects always contribute their entire endowment, which suggests that unconditional giving is a behavioral strategy that should be acknowledged. Many existing studies ignore or else fail to explain this presence of unconditional giving. For instance, Fischbacher and Gachter (2010) and Ambrus and Pathak (2011) focus only on free-rider and conditional cooperator types. We use the decisions in the experiments and a simple set of criteria to classify each subject in a given experimental match as either a *free-rider*, who always contributes zero; *altruist*, who contributes everything, unconditionally; or *conditional cooperator*, whose contribution depends on the contributions of others.

Second, we find that more than half of the subjects switch between these behavioral classifications across games with different information treatments. This switching behavior is consistent with our modification to the other-regarding preferences model, which links social comparisons to the outcomes and reference groups that are made salient by the information treatment. Overall, approximately 70% of the subjects who participated in the experiments behave in a manner consistent with the modified other-regarding preferences model.

Third, we show that learning is an important channel through which information and interaction structure affect contributions. After classifying the subjects as free-riders, altruists, or conditional cooperators, we compute the complete information Nash equilibrium of the one-shot game. This equilibrium provides a benchmark which we might expect players to converge to with repetition in the experiment, as they learn how to play the game. In fact, we find that in most cases, subjects' contributions do not converge to the corresponding Nash equilibrium. However, they are remarkably consistent with simulated behavior generated by the combined ORP-IEL model. By their classification, the free-riders and altruists converge almost immediately (or very quickly) to their

Nash equilibrium (and dominant strategy) contributions of zero and the endowment, respectively. The stability of their contribution decisions provides an anchor that accelerates learning by their neighbors in the network. As a result, conditional cooperators who are connected to free-riders or altruists learn faster than those who are not, and the degree of learning spreads contagiously around the circle network. This differential learning between free-riders/altruists and conditional cooperators, and the learning contagion effects are both observed in the experimental data and reproduced in the simulation data.

In the next section, we introduce the game and outline the behavioral model. Section 3.3 describes the experimental procedures and design. In Section 3.4, we derive the benchmark Nash equilibrium of the one-shot game and show that players may switch between behavioral strategies across games with different information treatments. In Section 3.5, we present the main findings from the data, and in Section 3.6, we present the learning results. Section 3.7 concludes and discusses avenues for further research.

3.2 The Network Public Goods Game

Let $I = \{1, ..., n\}$ be the set of players, each with an endowment $\omega_i = 1$. The players must each choose how much of their endowment to consume privately, and how much to contribute towards the public good. Player i's contribution to the public good is $y_i \in [0, 1]$ and $\mathbf{y} = (y_i)_{i \in I}$ is the profile of contributions made by all players.

The players are connected through a network, which is defined by an $n \times n$ matrix G, where $G(i, j) = 1$ if and only if player i is directly connected to player j. We assume $G(i, i) = 0$ for all $i \in I$. We use $N_i(G)$ to denote the set of direct neighbors for player i in the network G. Formally,

$$N_i(G) = \{j \in I \,|\, G(i, j) = 1\}. \tag{3.1}$$

Let $k_i(G) = |N_i(G)|$ be player i's degree.

The level of public good enjoyed by player i depends on his own contribution, the contributions made by his neighbors, and the marginal return to contributions, A, and is given by

$$Y^i(G,\mathbf{y}) = A \cdot \left[y_i + \sum_{j \in N_i(G)} y_j \right]. \tag{3.2}$$

Player i's payoff is

$$\pi_i(G,\mathbf{y}) = 1 - y_i + Y^i(G,\mathbf{y}). \tag{3.3}$$

We consider $A \in (0.5, 1)$ since it gives rise to a classical social dilemma. Within this range, an individual looking to maximize payoffs has a dominant strategy to free-ride, while the total welfare of all players is maximized by everyone contributing their entire endowment.

3.3 Experimental Design

The experiments were run in the Social Science Experimental Laboratory (SSEL) at the California Institute of Technology in August and September 2011. There were a total of 72 participants over 6 sessions. Each session lasted for approximately 1 hour and subjects earned an average payout of US\$25. Earnings in the experiment were denominated in tokens and converted into US dollars at the end of the experiment (500 tokens = US\$1). At the start of the experiment, the instructions were distributed and then read aloud, with summary slides projected at the front of the room. At the end of the instructions, the subjects participated in an unpaid 3-period practice match and answered four questions about the experiment and the way earnings were calculated.

There were 12 subjects in each session. A session was composed of 4 matches. At the start of each match, the 12 subjects were randomly divided into two groups of six players. In each group, the players were assigned to a position on the circle network shown in Figure 3.1 and then played 15 rounds of the voluntary contributions game described in Section 3.2.

In every round of a match, each subject received 100 tokens. During the round, we asked them how many tokens they wished to contribute towards some project and informed them that whatever

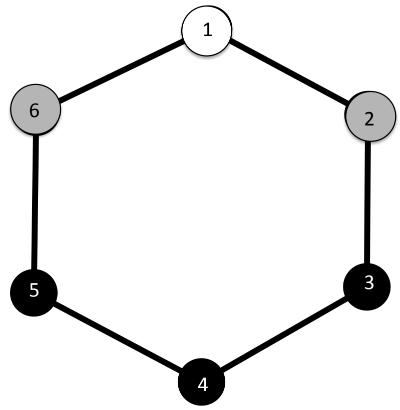


Figure 3.1: The circle network with 6 agents. As an example, the neighborhood for player 1 includes himself and his two neighbors, player 2 and player 6.

tokens they did not contribute, they could keep for themselves. We used a partners treatment, so that subjects stayed in the same group and at the same position in the network through every round of a given match. In four of the sessions, the 12 subjects were randomly rematched and repositioned before each match. In the other two sessions, subjects were randomly rematched and repositioned before match 1 and match 3 only. Their groups and positions were unchanged from match 1 to match 2 and from match 3 to match 4.

We calculated player payoffs as follows. First, we added the player's contribution to the sum of the contributions made by her direct neighbors. Then we multiplied the total by a return factor A (equal to 0.6 in all six experiments), and added this amount to the tokens that were kept by the player. That is, for player i, the payoff π_i in a particular round was given by

$$\pi_i = 100 - \text{ } i\text{'s contribution } + A \cdot (\text{ } i\text{'s contribution } + \sum \text{ } i\text{'s neighbors' contributions }).$$

In every round of a match, the player's location, her direct neighbors, and the return factor A were kept fixed. Moreover, the payoffs in one round did not depend on decisions made in previous rounds.

3.3.1 Information Treatments

After each round, the subjects were given the following summary information. In all matches, we showed them the amount they contributed to the project, the total contributions made in their neighborhood, and their payoff from the round. In addition, we reported one of the following information treatments.

(i) The average payoff received by **you and your direct neighbors**.

(ii) The average payoff received by **all six players in your group**.

(iii) The average contribution made by **you and your direct neighbors**.

(iv) The average contribution made by **all six players in your group**.

In a given match, all players faced the same information treatment, and we used the same information treatment in every round of a match. After the last player made his or her contribution decision, we displayed the round summary screen with all the information for approximately 8 to 10 seconds, after which the next round began. All reported information from previous rounds was displayed in a history panel at the bottom of the screen.[4] We ran the two groups simultaneously for each match so that subjects did not know which 5 out of the other 11 subjects they were matched with, let alone which two were their direct neighbors.

3.4 The Behavioral Model

In this section, we introduce the behavioral model, which combines a modified model of other-regarding preferences with the IEL model of Arifovic and Ledyard (2009).

3.4.1 Other-Regarding Preferences

The preferences in the model are an extension of the specification used by Arifovic and Ledyard (2009). Their model is consistent with the outcome-based or distributional models of Fehr & Schmidt

[4]Screenshots are provided in the Appendix along with the experiment instructions.

(1999) and Bolton & Ockenfels (2000). Each player evaluates outcomes with respect to a reference group of others, according to two underlying motivations. First, the player is altruistic, in that he cares about the well-being of the players in his reference group. Second, the player is concerned about fairness, in the sense that he dislikes outcomes in which he does not do as well as the others in his reference group. In Arifovic and Ledyard (2009), the reference group is just the set of all players in the game. Our innovation introduces two channels that link preferences to the information a player observes about others' decisions. The first channel is the player's reference group, which is induced by the *treatment group*, and the second is the information statistic, which he uses to evaluate outcomes.

Let R_i be player i's reference group, and assume that $i \in R_i$ and $|R_i| \geq 2$ for all $i \in I$. We consider two possible reference groups in this paper,

(1) the player's neighborhood, $R_i = N_i(G) \cup \{i\}$, and

(2) the whole set of players, $R_i = I$.

Let $\mathbf{S}_i$ be the reference statistic used by player i to evaluate outcomes. In this paper, we consider the case in which players evaluate outcomes using either

(1) average contribution in their reference group, $\mathbf{S}_i = \sum_{j \in R_i} \frac{y_j}{|R_i|}$, or

(2) average payoff in their reference group, $\mathbf{S}_i = \sum_{j \in R_i} \frac{\pi_j}{|R_i|}$.

Each player has an underlying type, which consists of an altruism parameter, $\beta^i \geq 0$ and an envy parameter, $\gamma^i \geq 0$. We assume this underlying type is stable and independent of both the reference group and reference point. Thus, we can write player i's utility function as

$$U_i(G, \mathbf{y}) = \pi_i(G, \mathbf{y}) + \beta^i \cdot \mathbf{S}_i - \gamma^i \cdot \max\left\{0\,,\, \Delta\mathbf{S}_i\right\}, \tag{3.4}$$

where $\Delta\mathbf{S}_i$ takes one of two forms. If player i's reference point is the average contribution in her reference group, then

$$\Delta\mathbf{S}_i = y_i - \sum_{j \in R_i} \frac{y_j}{|R_i|}. \tag{3.5}$$

On the other hand, if player i's reference point is the average payoff in her reference group, then

$$\Delta\mathbf{S}_i = \sum_{j \in R_i} \frac{\pi_j}{|R_i|} - \pi_i. \tag{3.6}$$

3.4.2 Equilibrium Behavior

Given a set of types, reference groups, and reference statistics, it is relatively straightforward to solve the (complete information) Nash equilibrium of the game. First, consider the case in which the reference statistic is *average contribution* for all players.

Proposition 3.4.1. *Suppose* $\{(\beta^i, \gamma^i)\}_{i \in I}$ *are the players' preference parameters,* $\{R_i\}_{i \in I}$ *are their reference groups, and* $\mathbf{S}_i = \sum_{j \in R_i} \frac{y_j}{|R_i|}$ *are their reference statistics. Then there exists a Nash equilibrium* $\mathbf{y}^*$*, in which each player* i *chooses a contribution equal to*

$$(a) \quad y_i^* = 0 \qquad \text{if} \quad \beta^i \le (1-A)|R_i| \tag{3.7}$$

$$(b) \quad y_i^* = \sum_{j \in R_i} \frac{y_j^*}{|R_i|} \qquad \text{if} \quad (1-A)|R_i| < \beta^i < (1-A+\gamma^i)|R_i| - \gamma^i \tag{3.8}$$

$$(c) \quad y_i^* = 1 \qquad \text{if} \quad \beta^i \ge (1-A+\gamma^i)|R_i| - \gamma^i. \tag{3.9}$$

If there exists some player j *who satisfies either of the conditions in (a) or (b), then the equilibrium is unique. Otherwise, there are a continuum of equilibria in which* $y_i^* = y_j^*$ *for all players* $i, j \in I$.

Proof 6. *See Appendix.*

Proposition 3.4.1 says that in equilibrium, a player either contributes everything, contributes nothing, or contributes the average contribution made by the players in her reference group. The conditions on β^i and γ^i have a natural interpretation. The inequality in (3.7) says that if player

i's altruistic parameter, β^i is sufficiently low, then his best response to any profile of others' contributions is to contribute zero. In this case, we refer to him as a *free-rider.* On the other hand, the inequality in (3.9) implies that if player i's altruistic parameter is high enough relative to his fairness parameter γ^i, then his best response is always to contribute everything, making him an *altruist.* For a player who falls into case (b), for any given profile of others' contributions, his best response is to contribute the average amount contributed by the players in his reference group. We refer to such a player as a *conditional cooperator.*

Deriving the Nash equilibrium of the game for a given realization of types, and reference groups, when the reference statistic is average contribution, involves a few simple steps. First, we classify the altruists and free-riders, whose equilibrium behavior is unconditional. Then we find the equilibrium contributions of the conditional cooperators by solving a system of linear equalities. As long as there is at least one free-rider or altruist, there is a unique solution to this system. However, if all players are conditional cooperators, then the system has a continuum of solutions in which all players contribute the same amount.

Next, suppose that the reference statistic is *average payoff* for all players.

Proposition 3.4.2. *Suppose* $\{(\beta^i, \gamma^i)\}_{i\in I}$ *are the players' preference parameters,* $\{R_i\}_{i\in I}$ *are their reference groups, and* $\mathbf{S}_i = \sum\limits_{j\in R_i} \frac{\pi_j}{|R_i|}$ *are their reference statistics.. Then there exists an equilibrium in which player* i *chooses a contribution equal to*

$$(a) \qquad y_i^* = 0 \qquad \text{if} \quad \beta^i \leq \frac{(1-A)|R_i|}{A\Big(|N_i(G)\cap R_i|+1\Big)-1} \tag{3.10}$$

$$(b) \qquad y_i^* = 1 \qquad \text{if} \quad \beta^i \geq \gamma^i + \frac{(1-A)(1+\gamma^i)|R_i|}{A\Big(|N_i(G)\cap R_i|+1\Big)-1}. \tag{3.11}$$

If player i *does not satisfy either of the conditions in (a) or (b), her best response to a profile of others' contributions is to choose the feasible contribution that minimizes the absolute difference between her own payoff and the average payoff in her reference group.*

Proof 7. *See Appendix.*

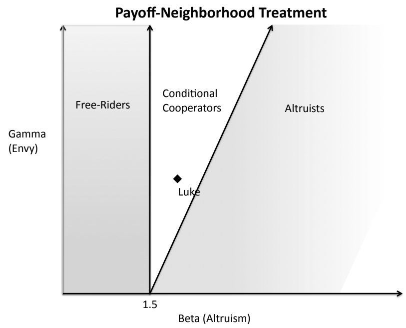
Proposition 3.4.2 provides similar predictions as Proposition 3.4.1. A player who is sufficiently self-interested will free-ride, while a player who is sufficiently altruistic, relative to his envy, will give everything. The equilibrium behavior of a conditional cooperator is slightly different, due to feasibility constraints. While a conditional cooperator would like to equalize his payoff with the average payoff in his reference group, this will not always be possible, since he cannot contribute less than zero or more than his endowment. As a result, the solution to the set of simultaneous inequalities may be on the boundary.

These propositions imply that, even in equilibrium, a player with fixed type parameters β^i and γ^i, may behave differently when their reference group and reference statistic change. This is because the critical cutoffs in the type space that define free-riders, altruists, and conditional cooperators depend on the reference group and reference statistic. Given our assumption that players use the treatment group as their reference group and the treatment statistic as the reference statistic, our model predicts variation in behavior as we change the information treatment. The sequence in Figure 3.2 illustrates how a player with fixed type parameters faces different incentives under the four alternative treatments used in the experiments.

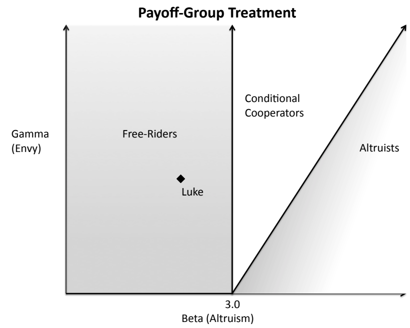
3.4.3 Learning

While the other-regarding preferences (ORP) are crucial for explaining behavior, they are still not sufficient to explain the full pattern of contributions in the repeated game. Our full behavioral model combines the modified ORP model with the IEL theory of learning developed by Arifovic & Ledyard (2009).

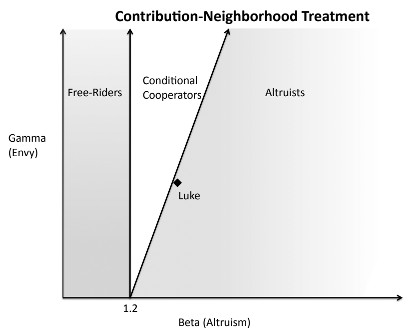
The main features of the IEL model are described by the following. Each player i keeps a finite set of remembered actions in period t, denoted by $A_t^i \subset [0, 100]$. Suppose that $|A_t^i| = J$ for all i and all t. Furthermore, each player uses a forgone utility function, $u^i(a_{j,t}^i | r^i(\mathbf{y}_t))$, which specifies the utility that player i thinks he could have earned by playing alternative $a_{j,t}^i$ in period t, given the information $r^i(\mathbf{y}_t)$ that he received at the end of period t. The three stages of the model are



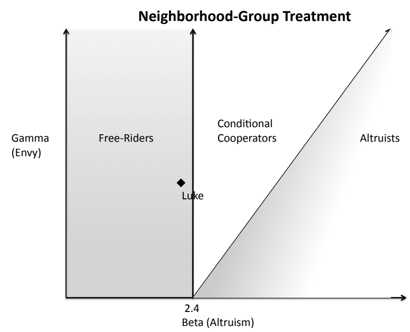
(a)



(b)



(c)



(d)

Figure 3.2: Switching behavior predicted by the modified other-regarding preferences model. The player "Luke" has fixed type parameters (β^i, γ^i), while the cutoff regions change with the information treatment.

1. *Experimentation.* For each player, on each slot $j = 1, ..., J$ of their remembered set, experimentation occurs with probability ρ. If experimentation occurs, then the old alternative in slot j, $a^i_{j,t}$ is replaced by a new action, which is randomly selected from $[0, 100]$ according to a normal distribution with mean $a^i_{j,t}$ and standard deviation σ.

2. *Replication.* For each player, and each slot $j = 1, ..., J$, the alternative filling slot j in period $t + 1$ is chosen as follows. Pick two member of A^i_t randomly with replacement, with uniform probability. Call these alternatives $a^i_{k,t}$ and $a^i_{l,t}$. Then set

$$a^i_{j,t+1} = \begin{cases} a^i_{k,t} & \text{if } u^i(a^i_{k,t}|r^i(x_t)) \geq u^i(a^i_{l,t}|r^i(x_t)) \\ a^i_{l,t} & \text{if } u^i(a^i_{k,t}|r^i(x_t)) < u^i(a^i_{l,t}|r^i(x_t)). \end{cases} \tag{3.12}$$

This procedure favors alternatives with a lot of replicates in the remembered set at period t, and alternatives that would have paid well in period t, had they been used.

3. *Selection.* In period $t+1$, alternative $a^i_{k,t+1} \in A^i_{t+1}$ is selected with probability proportional to the forgone utility from that alternative, given by

$$\mu^i_{k,t+1} = \frac{u^i(a^i_{k,t+1}|r^i(x_t)) - \varepsilon^i_{t+1}}{\sum_{j=1}^{J}\left[u^i(a^i_{j,t+1}|r^i(x_t)) - \varepsilon^i_{t+1}\right]} \tag{3.13}$$

for all $i \in \{1, ..., N\}$ and all $k \in \{1, ..., J\}$, where

$$\varepsilon^i_{t+1} = \min_{a \in A^i_{t+1}} \left\{0, u^i(a|r^i(x_t))\right\}.$$

We assume that the process is initialized by randomly populating A^i_1 with J uniform draws from $[0, 100]$, and setting $\mu^i_{k,1} = \frac{1}{J}$ for all slots k. Given the design of our experiments, we assume that the forgone utility function depends on the information treatment and the network structure. We consider each treatment separately.

Treatment 1 — Average Payoff in Neighborhood

In this case, the players observe their own contribution, y^i_t, their own payoff, π_i and the average payoff in their neighborhood, which is given by

$$\bar{\pi}_i = \frac{\pi_i}{|N_i(G) \cup \{i\}|} + \sum_{j \in N_i(G)} \frac{\pi_j}{|N_i(G) \cup \{i\}|}.$$

Furthermore, $|N_i(G) \cup \{i\}|$ and G are known by each subject. Thus, if player i chooses y^i_t in period t, her own perceived payoff from choosing a^i instead of y^i_t, given what her neighbors did would have been

$$\hat{\pi}_i = \pi_i - (a^i - y^i_t)(1 - A),$$

while the average payoff of her reference group (the players in her neighborhood, including herself) would have been

$$\hat{\bar{\pi}}_i = \bar{\pi}_i + (a^i - y^i_t)\left[\frac{|N_i(G)|}{|N_i(G) \cup \{i\}|}A - \frac{(1-A)}{|N_i(G) \cup \{i\}|}\right].$$

From these two counterfactuals, we can write the forgone utility function for player i as

$$\begin{aligned} u^i(a^i|y^i_t, \pi_i, \bar{\pi}_i) &= \pi_i - (1-A)(a^i - y^i_t) \\ &+ \beta^i(\bar{\pi}_i + (a^i - y^i_t)\left(\frac{|N_i(G)|}{|N_i(G) \cup \{i\}|}A - \frac{(1-A)}{|N_i(G) \cup \{i\}}\right) \\ &- \gamma^i \cdot \max\left\{0, \bar{\pi}_i - \pi_i + (a^i - y^i_t)\frac{|N_i(G)|}{|N_i(G) \cup \{i\}|}\right\}. \end{aligned}$$

Notice that the counterfactuals are all assessed under the assumption that other players' decisions are kept fixed.

Treatment 2 — Average Payoff in Group

In Treatment 2, subjects observe the average payoff of all the players in their group. Thus, they see their own contribution y^i_t, their own payoff π_i, and the group average payoff given by

$$\bar{\pi} = \frac{\pi_i}{n} + \sum_{j \in I} \frac{\pi_j}{n}.$$

In such a case, the counterfactual own payoff and group average payoff allow us to write the forgone utility function for player i from choosing a contribution a^i instead of y^i_t as

$$\begin{aligned} u^i(a^i|y^i_t, \pi_i, \bar{\pi}) = \pi_i - (1-A)(a^i - y^i_t) + \beta^i\left(\bar{\pi} + (a^i - y^i_t)\big(\frac{|N_i(G)|A - (1-A)}{n}\big)\right) \\ -\gamma^i \cdot \max\left\{0, \bar{\pi} + \frac{(a^i - y^i_t)\big(|N_i(G)|A - (N+1)(1-A)\big)}{n}\right\}. \end{aligned}$$

Treatment 3 — Average Contribution in Neighborhood

For Treatment 3, subjects observe their own contribution, y_t^i, their own payoff, π_i and the average contribution made in their neighborhood,

$$\bar{y}_i = \sum_{j \in N_i(G) \cup \{i\}} \frac{y_t^j}{|N_i(G) \cup \{i\}|}.$$

Determining the counterfactual own payoff, and the counterfactual average neighborhood contribution, we can write the forgone utility function for this case as

$$\begin{aligned} u^i(a^i | y_t^i, \pi_i, \bar{y}_i) = \pi_i - (1-A)(a^i - y_t^i) + \beta^i \left(\bar{y}_i + \frac{(a^i - y_t^i)}{|N_i(G) \cup \{i\}|} \right) \\ -\gamma^i \cdot \max \left\{ 0, a^i - \bar{y}_i - \frac{(a^i - y_t^i)}{|N_i(G) \cup \{i\}|} \right\}. \end{aligned}$$

Treatment 4 — Average Contribution in Group

Finally, for Treatment 4, subjects observe own contribution, y_t^i, own payoff, π_i and the group average contribution,

$$\bar{y} = \sum_{j \in I} \frac{y_t^j}{n}.$$

In this case, the relevant counterfactuals, assuming others' decisions do not change, generate the forgone utility function given by

$$\begin{aligned} u^i(a^i | y_t^i, \pi_i, \bar{y}) = \pi_i - (1-A)(a^i - y_t^i) + \beta^i \left(\bar{y} + \frac{(a^i - y_t^i)}{n} \right) \\ -\gamma^i \cdot \max \left\{ 0, a^i - \bar{y} - \frac{(a^i - y_t^i)}{n} \right\}. \end{aligned}$$

In Section 3.6, we return to the learning model, and describe the simulation procedures and the learning results. In the next section, we turn instead to the experimental findings.

3.5 Experimental Findings

We begin by describing the aggregate data, which paints a picture consistent with previous experimental findings. As Figure 3.3 shows, the average contribution exhibits a steady decay from just over half the endowment in the first round, to about 20% of the endowment in the last round. This pattern of decay is also consistent in each of the four information treatments.

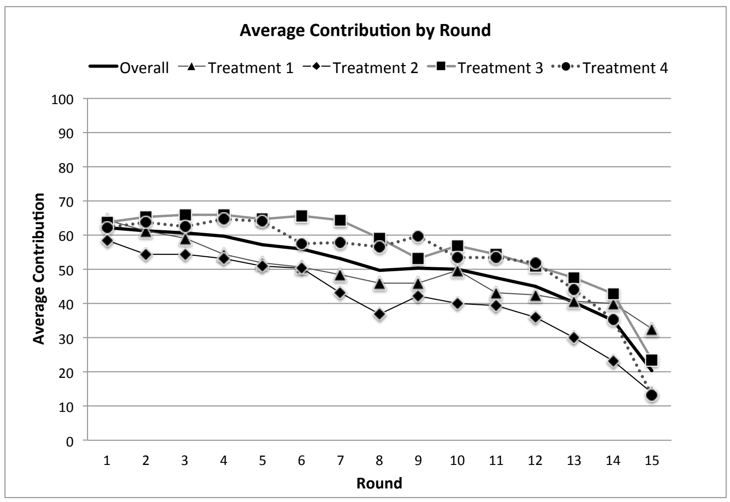


Figure 3.3: Average contribution over all matches by round. Treatment 1 = ave. payoff in neighborhood, Treatment 2 = ave. payoff in group, Treatment 3 = ave. contribution in neighborhood, Treatment 4 = ave. contribution in group

Between treatments, average contributions are slightly higher in Treatments 3 and 4, when players observe the average contributions of others, than in Treatments 1 and 2, when they observe the average payoffs of others. This may be a result of more free-riders in the average payoff treatments than the average contribution treatments, which is consistent with the higher cutoffs for free-riding predicted by the model. On the other hand, there is little evidence that average contributions differ depending on whether the subjects are shown information about the whole group or just about their neighbors. Nevertheless, there is much more variation when we turn to look at the individual match level data.

3.5.1 Subject Behavior

In order to analyze the data at the individual match level, I first classify the subjects by their behavior as a free-rider, altruist, or conditional cooperator. I use a simple set of criteria to classify the subjects, based on their choices in the penultimate three rounds of the matches. The criteria are as follows.

Consider the subject's contribution decisions in Rounds 12, 13, and 14.

(1) If the subject's contribution is less than 10 in at least 2 of the 3 rounds, she is classified as a free-rider.

(2) If the subject's contribution is greater than 90 in at least 2 of the 3 rounds, she is classified as an altruist.

(3) If the subject's contribution does not satisfy either of the conditions in (1) or (2), she is classified as a conditional cooperator.

In the data, subjects' decisions exhibit all three types of behavior predicted by the modified other-regarding preferences model. Table 1 summarizes the frequency of each behavior and highlights the considerable heterogeneity among the subject pool. Furthermore, there are a large variety of configurations of these classes in the 48 Matches. Table 2 shows the frequency of 6 broad categories of configurations. Within each of the reported categories there is even more variation regarding, for instance, the number of free-riders and conditional cooperators, whether free-riders are adjacent or separated, or whether or not a free-rider and altruist are adjacent.

Table 3.1: Player Behavior Classification

Behavior type	Frequency
Free-Rider (F)	74
Altruist (A)	53
Conditional Cooperator (C)	161
Total	288

Table 3.2: Configuration of Players: F = Free-rider, A = Altruist, C = Conditional Cooperator

Configuration description	Frequency
F & C only	18
A & C only	10
F only	1
A only	2
C only	5
F, A, & C	12
Total	48

In order to check how well my classification fits with the data, I examine the average contribution made by each class of player in each round. Figure 3.4 shows the average contribution by round for each class. The figure illustrates two key points. First, the free-riders take time to converge to free-riding. In contrast, the altruists tend to converge very quickly to full contributions. Second, the classification procedure provides a fairly robust fit to the decisions of the free-riders and altruists, even in the earlier rounds which are not used for classification. Out of the 74 free-riders, 15 contribute 0 in every round, and 29 of them have an average contribution (over all rounds) less than 5. Similarly, out of the 53 altruists, 18 contribute 100 in every round, and 41 of them have an average contribution (over all rounds) greater than 90. Thus, there are some classified free-riders who contribute nothing from the beginning, and almost all classified altruists contribute very close to their entire endowment through every round of the game.

In Figure 3.5, I also show the average contribution for the conditional cooperators, when they are broken down by the classification of their direct neighbors. For instance, notice that the average contribution for conditional cooperators with two altruist neighbors is close to full contribution, while the average contribution for conditional cooperators with two free-rider neighbors is decreasing to zero (there are only 2 players in this situation). For the other subclasses of conditional cooperators, the average contributions are all in line with what we should expect. Those with an altruist neighbor contribute higher (on average) than those with a free-rider neighbor. For the most part, the average contribution in the subclasses also appears to decline over time.

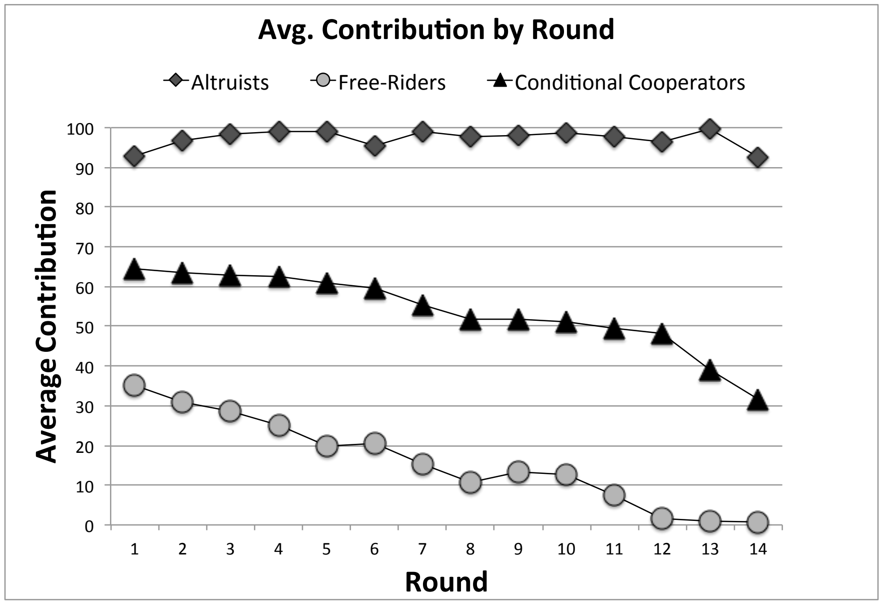


Figure 3.4: Average contribution by round for each class

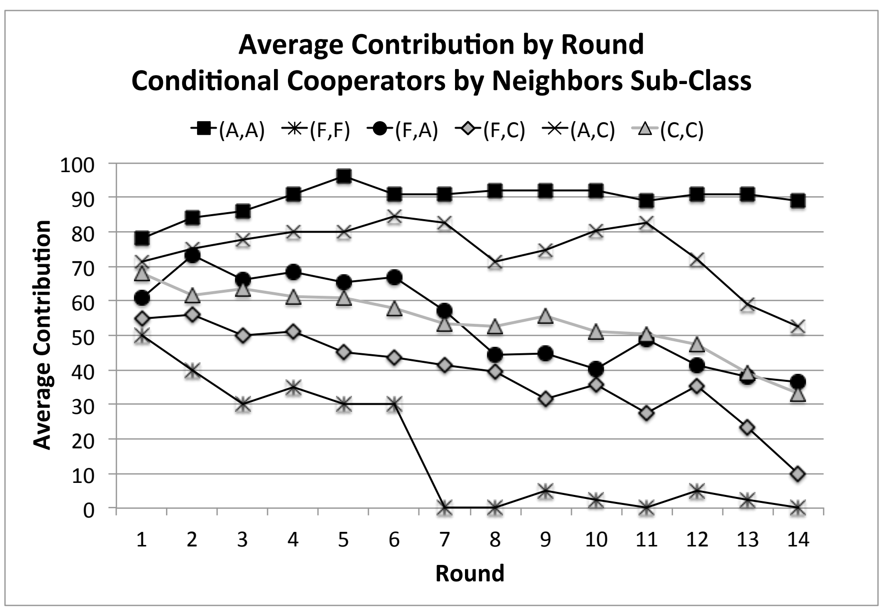


Figure 3.5: Average contribution by round for each subclass of conditional cooperator

3.5.1.1 Switching Behavior in the Experiments

In the data there are many subjects who have different classifications in different matches. Using the criteria described above, there are 37 subjects (51%) who are classified differently in different matches. Of these 37 subjects, 15 exhibit switching that is perfectly consistent with the restrictions of the theory. There may be several reasons why the other subjects who switch do not appear to do so in a manner consistent with the theory. For one thing, subjects may be making mistakes or still learning about their incentives in the game, particularly in the first or second matches that they play. However, this is a difficult phenomenon to identify or allow for in the analysis. Alternatively, there may be some incorrect classification by the simple criteria that I use in this paper.[5] All in all, the analysis in this section suggests that, out of the 72 subjects who participated in these experiments, 50 of them (70%) exhibit revealed behavior that can be explained by the modified other-regarding preferences model.

Table 3.3: Switching Behavior by Subject

Subject's Classifications	Number of Subjects
Free-Rider (F) in all 4 matches	8
Conditional Cooperator (CC) in all 4 matches	24
Altruist (A) in all 4 matches	3
Either F or C in each match	15
Either C or A in each match	14
Either F or A in each match	2
Each of F, C & A in some match	6
Total	72

3.5.2 Comparison to the Benchmark Nash Equilibrium

In Section 3.4, we derived the conditions for solving for the complete information Nash equilibrium of the stage game, assuming players have other-regarding preferences. Except for the case when all six players are conditional cooperators (CCs), there is a unique Nash equilibrium for a given configuration of free-riders, altruists, and CCs. Even though subjects do not have complete information

[5] One way to address this is to try to elicit players' revealed types from a separate social dilemma type game that might allow us to parametrize the subjects' types before they play the repeated public goods games. An approach similar to the one used by Fischbacher et al. (2001) in which players play submit their contribution decisions via the strategy method is one alternative.

in the experiments, a natural hypothesis is that players' contributions will converge fairly close to the Nash equilibrium after 10 or 12 rounds of playing with the same group. In fact, we show that players' contributions do not always converge to the corresponding Nash equilibrium.

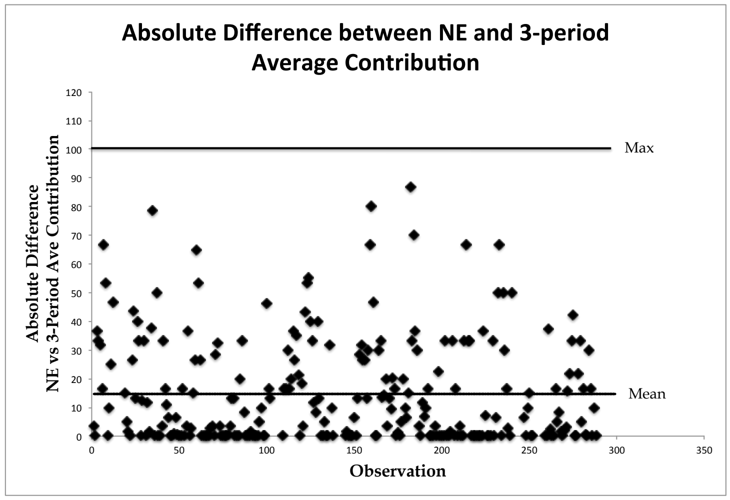
Since each match is unique, with possibly a different configuration of types, we focus on the absolute differences between the Nash equilibrium contributions, and the 3-period average contribution from rounds 12–14. Panel (a) of Figure 3.6 is a scatterplot of the differences for each player in all matches. It shows that there are many points scattered far away from zero. The variance is significantly larger than we need to conclude that contributions converge to the Nash equilibrium. The mean difference is 14.69, while the standard deviation is 18.48, almost one fifth of the endowment. This picture is also slightly misleading, since any free-rider or altruist types are so classified based on their contributions in the penultimate three rounds of the game. Thus, including them in the calculation biases the average difference between Nash equilibrium and the 3-period average contribution towards zero.

In panel (b) of Figure 3.6, we remove the altruist and free-rider types. In this case, the picture is even more bleak. The mean absolute difference increases to 26.19, while the standard deviation stays fairly constant at 18.64. This is hardly overwhelming evidence in support of the hypothesis that subjects converge to the corresponding Nash equilibrium given the underlying types. Such disparity between the data and the Nash equilibrium contributions suggests that if conditionally cooperative subjects are learning how to play the game over time, that learning is incomplete.

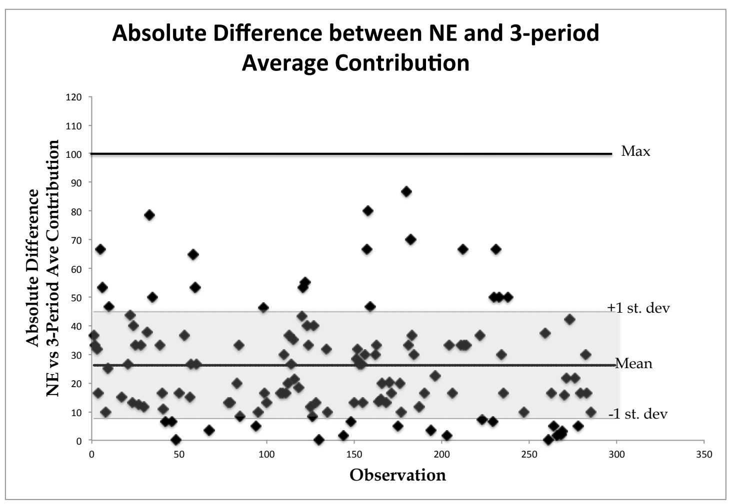
From the individual match level data, we find evidence for the following two results.

Result 3.5.1. *Conditional cooperators' contributions come closer to the corresponding Nash equilibrium contribution when there are more free-riders or altruists in the group.*

This first result suggests that the composition of players in the match affects the level of contributions made by conditional cooperators. Below, we argue that this result is driven by differential learning. Free riders and altruist learn to play their dominant strategies (zero and full contributions, respectively) relatively fast in the repeated game. The stability of their contributions provides an anchor which enhances learning by their conditional cooperator neighbors. The more anchors in the



(a) All player types. 258 observations, mean absolute difference = 14.69, std. error = 18.48



(b) Conditional cooperators only. 132 observations, mean absolute difference = 26.19, std. error = 18.64

Figure 3.6: Absolute difference between 3-period ave. contribution (Rounds 12–14) and the corresponding Nash equilibrium. These figures suggest that contributions do not converge to the Nash equilibrium, even on average.

group, the faster the conditional cooperators learn, and the closer they come to playing the Nash equilibrium by the end of the game.

The second result emphasizes the importance of the players' relative positions in the network, given the group's composition. In particular, players that are positioned next to a free-rider or next to an altruist tend to come closer to contributing the corresponding Nash equilibrium level.

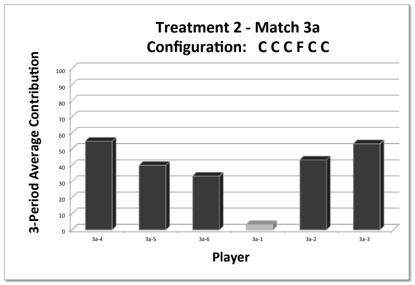
Result 3.5.2. *Individual contributions in a given match tend to be closer to (further from) the benchmark Nash equilibrium for conditional cooperators who are positioned closer to (further from) the free-rider or altruists in the group.*

The intuition for this result is that, since free-riders and altruists provide an anchor for learning by their conditional cooperator neighbors, learning spreads contagiously through the network. For instance, if the group consists of a single free-rider and five conditional cooperators, the conditional cooperators who are next to the free-rider will learn faster and converge closer to the corresponding Nash equilibrium of zero contributions. The learning spreads more slowly to their other neighbors, and slowest to the conditional cooperator who is farthest removed from the free-rider in the network.

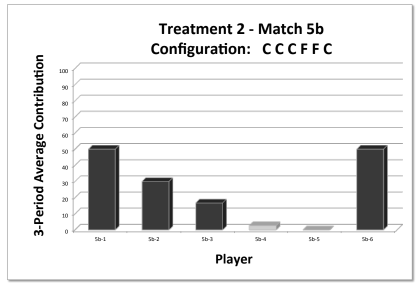
There is evidence supporting both of these results across each of the four information treatments. A full analysis is provided in Appendix B. There, we summarize the evidence from the individual matches broken down by the information treatment. However, we report several representative cases in this section, in order to highlight the results. First consider Result 3.5.1. In Figure 3.7, we present the actual 3-period average contributions in three matches played under Treatment 2, across various sessions. For each match, the player number is shown on the horizontal axis, with each player next to his neighbors. Note that the two end players are connected to each other. Where possible, we have arranged the order to show any free-riders or altruists in the middle.

In each of the three matches in Figure 3.7, the composition of types includes only free-riders and conditional cooperators, which implies that the benchmark Nash equilibrium for each player is to contribute zero. Moving from left to right, match 3a includes one free-rider, match 5b includes two free-riders, and match 2a includes three free-riders. As demonstrated by the figures, the average contributions made by the conditional cooperator types are declining as we move from the left panel (fewer free-riders) to the right panel (more free-riders). The result is also captured under Treatment 4, by matches 5b, 1a, and 4b, which consist of conditional cooperators grouped with two, three, and four altruists, respectively. In these cases, the corresponding Nash equilibrium contributions are 100 for all subjects. Moving from left to right, Figure 3.8 indicates that the average contribution of the conditional cooperators is increasing in the number of altruists.

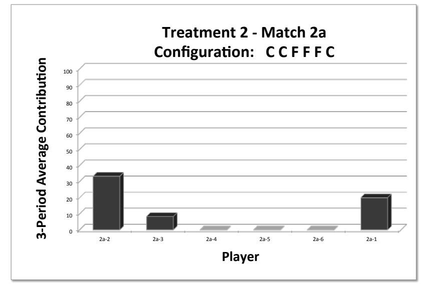
Next, consider Result 3.5.2. First of all, this result is also demonstrated remarkably well by the matches shown in Figure 3.7. In each of the three matches, contributions are increasing as we move in either direction, away from the free-riders. The result is also highlighted in Figure 3.9, which reproduces match 1b under Treatment 1, match 4a under Treatment 2, and matches 1a, 2b, and 3a, under Treatment 3.



(a) Match 3a

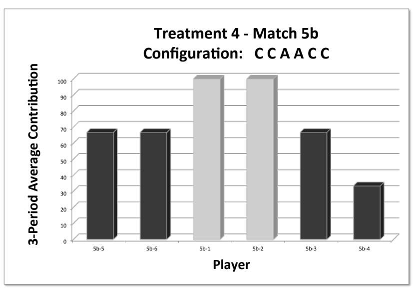


(b) Match 5b

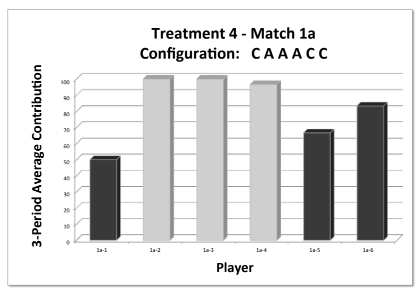


(c) Match 2a

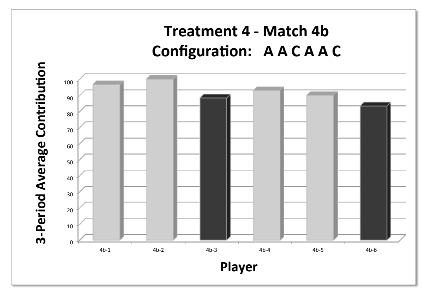
Figure 3.7: Treatment 2 matches with free-riders and CCs only. Free-riders' contributions are shown in gray, while the contributions of conditional cooperators are shown in black.



(a) Match 5b

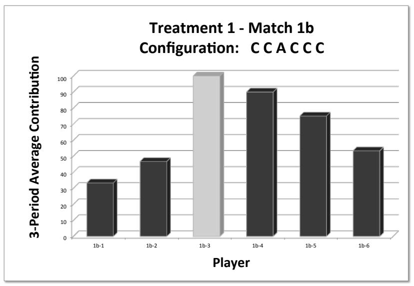


(b) Match 1a

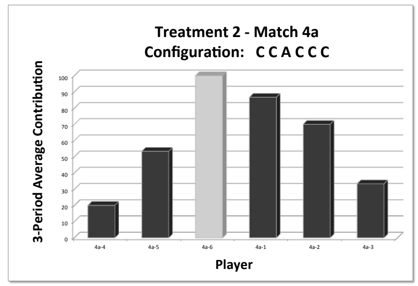


(c) Match 4b

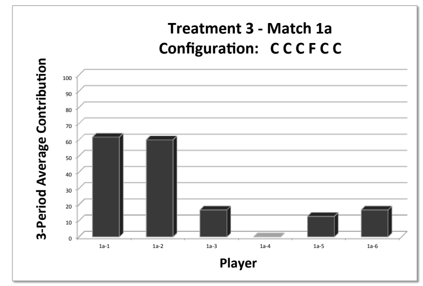
Figure 3.8: Treatment 4 matches with altruists and CCs only. Altruists' contributions are shown in gray, while the contributions of conditional cooperators are shown in black.



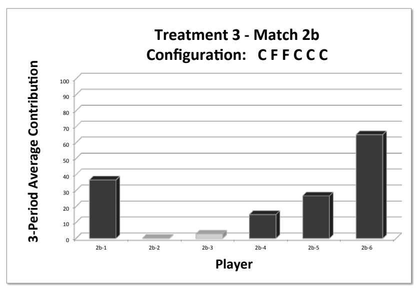
(a) Treatment 1 Match 1b



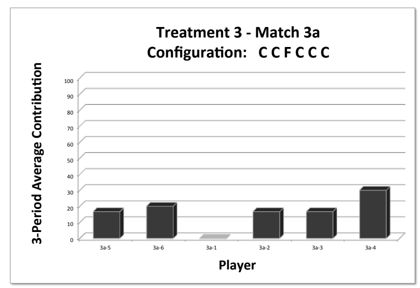
(b) Treatment 2 Match 4a



(c) Treatment 3 Match 1a



(d) Treatment 3 Match 2b



(e) Treatment 3 Match 3a

Figure 3.9: Effect of players' positions on contributions. Gray bars indicate either a free-rider or altruist. Conditional cooperators are shown in black.

3.6 Learning Results

In order to explain Result 3.5.1 and Result 3.5.2, we conduct learning simulations using a combined model of Individual Evolutionary Learning (IEL) and the modified other-regarding preferences model. In this section, we describe the simulations and present the learning results. These results show that the nonconvergence to the benchmark Nash equilibrium levels and the two features of the data described in Result 3.5.1 and Result 3.5.2 are all consistent with the combined ORP-learning model.

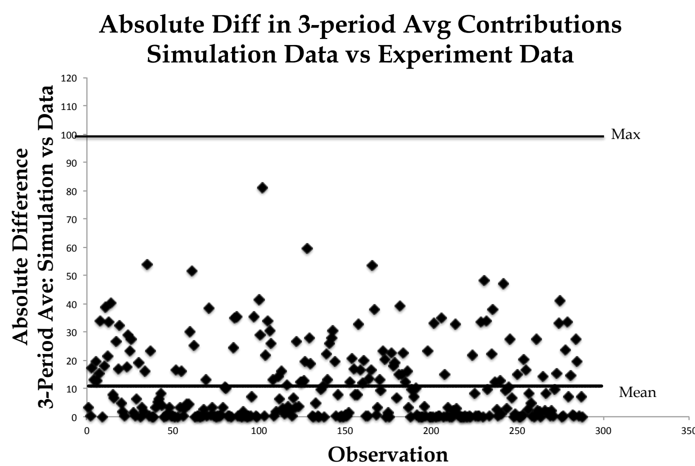
3.6.1 Simulations

For the simulations, we used the following parameter values to calibrate IEL. The size of the remembered set for each player is set to $J = 100$ and the probability of experimentation $\rho = 0.033$. The new alternative is selected according to a normal distribution, conditional on the set of possible contributions $[0, 100]$ with mean equal to the old alternative and a standard deviation equal to $\sigma = 10$. These are the same values of J, ρ, and σ used by Arifovic & Ledyard (2009), who show that their results are not sensitive to the precise values of these parameters. Likewise, in our simulations, the particular values of these parameters do not induce any significant changes in the results.

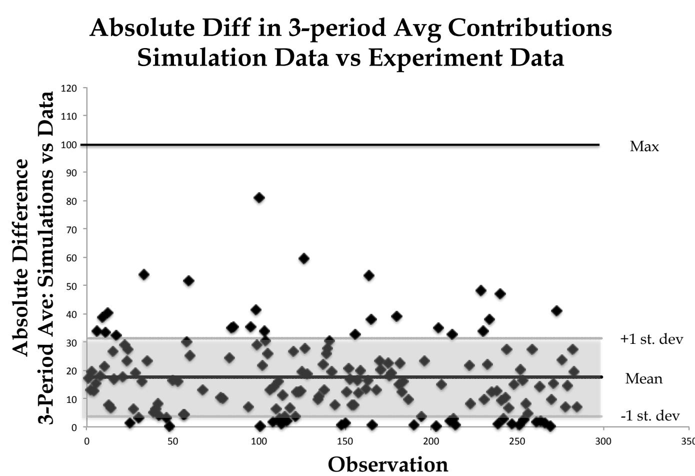
We ran 250 iterations for each simulation, and a simulation for each individual match in the dataset, tailored to the classified configuration of players, and the information treatment used for the match. For each information treatment, the cutoffs that characterize the free-riders and altruists are slightly different. Thus, in order to ensure that the configuration of players is consistent in the simulation with the corresponding individual match, we assign the free-riders and altruists preference parameters that satisfy the appropriate cutoff for the corresponding information treatment. For the conditional cooperators, each iteration of the simulation draws a new pair of preference parameters from a range that only permits incentives for conditional cooperation. This ensures that, in each iteration of each simulation, the actual configuration of players, as classified according to the criteria discussed above, remains the same.

3.6.2 Simulation Results

The scatterplots in Figure 3.10 show the absolute differences between the 3-period average contribution (from Rounds 12–14) in the simulations and the data. Panel (a) shows all players while panel (b) shows only the conditional cooperators in order to remove the downward bias from classification of the free-riders and altruists. The absolute differences are, on average, significantly improved over the differences between the data and the Nash equilibrium. Still, the variance is not particularly small for either of these cases. Overall, the simulations reduce the mean absolute difference by 26% from 14.69 to 10.85, using all players, and by 33% from 26.19 to 17.48, using just the conditional cooperators.



(a) All players



(b) Conditional cooperators only

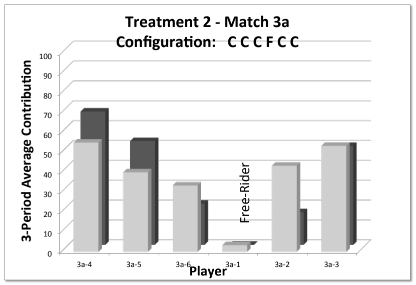
Figure 3.10: Differences between 3-period ave. contribution: Simulations vs data

The consistency between the simulations and data is much more evident at the individual match level. In particular, the learning simulations reproduce the findings summarized by Result 3.5.1 and Result 3.5.2. Here, we provide a summary which compares the simulation results with the evidence of the two main results discussed in Section 3.5. In each figure below, the 6 players in a match are shown together, next to their neighbors on the horizontal axis, with the exception of the players on the ends who are also connected to each other. On the vertical axis, we plot the simulation average in blue (behind) and the data average in red (in front). The main features of the data are reproduced in the learning simulations for a remarkably high percentage of the matches.

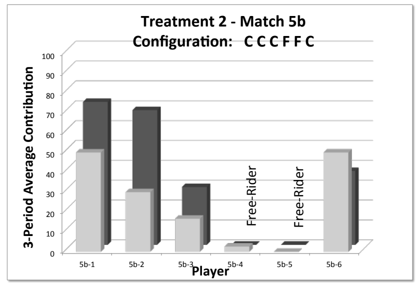
First consider Result 3.5.1. In Figure 3.11, we show that the simulations (blue) replicate the data (red) quite closely. From left to right, match 3a includes one free-rider, match 5b includes two free-riders, and match 2a includes three free-riders. As in the data, the simulations also suggest that the average contributions made by the conditional cooperators are declining as we move from the left panel across to the right panel.

The same is true for matches 5b, 1a, and 4b under Treatment 4, shown in Figure 3.12. In these matches, which consist of two, three, and four altruists, respectively, both the simulations and the data show that the average contribution of the conditional cooperators is increasing in the number of altruists.

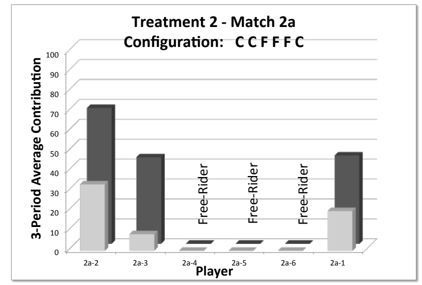
Next, consider Result 3.5.2. As discussed in Section 3.5, this result is also supported by the matches in Figure 3.11. In each of the three matches, both the simulated and actual contributions get noticeably higher for players as we move away from the free-riders. The simulations are also consistent with Result 3.5.2 in the matches displayed in Figure 3.13.



(a) Match 3a

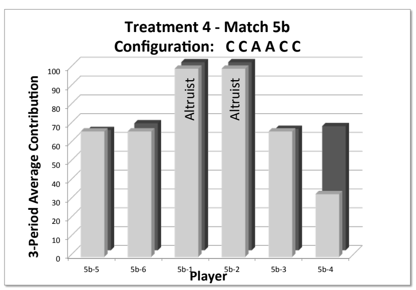


(b) Match 5b

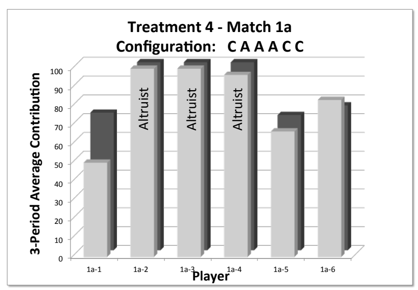


(c) Match 2a

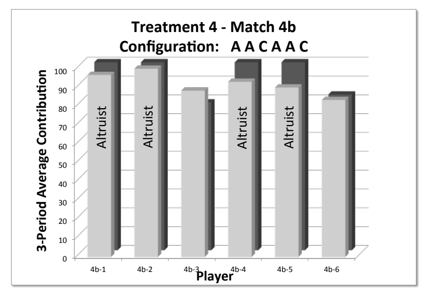
Figure 3.11: Treatment 2 matches with free-riders and CCs only. Simulation averages are shown in blue (behind), while the data averages are shown in red (front).



(a) Match 5b

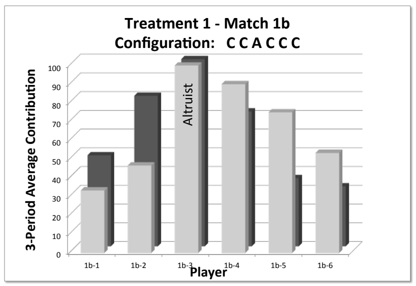


(b) Match 1a

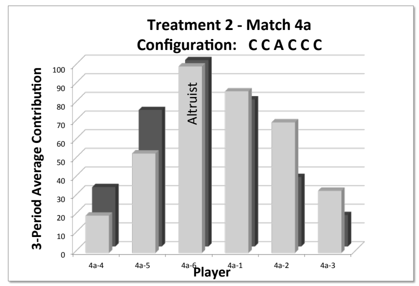


(c) Match 4b

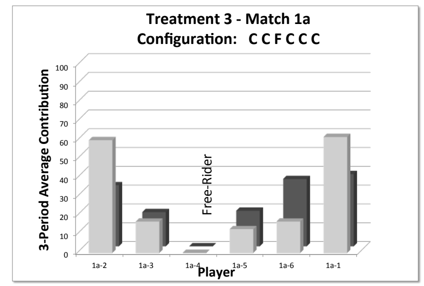
Figure 3.12: Treatment 4 matches with altruists and CCs only. Simulation averages are shown in blue (behind), while the data averages are shown in red (front).



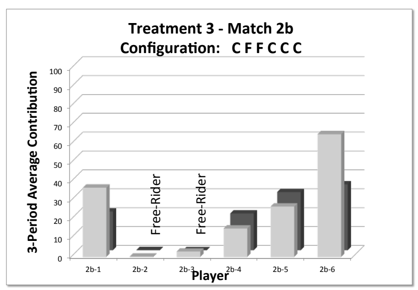
(a) Treatment 1 Match 1b



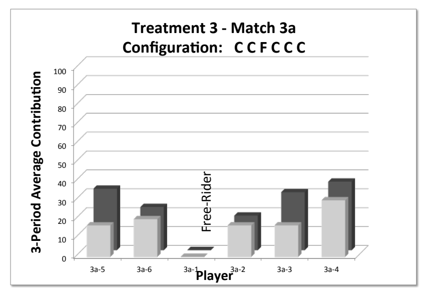
(b) Treatment 2 Match 4a



(c) Treatment 3 Match 1a



(d) Treatment 3 Match 2b



(e) Treatment 3 Match 3a

Figure 3.13: Learning contagion. Simulation averages are shown in blue (behind), while the data averages are shown in red (front).

3.7 Conclusion

In this paper, we find evidence that both the information shown to players between rounds and the neighborhood structure of interaction affect voluntary contributions to a public good on a network. Using a series of experiments, we show that contributions are consistent with a model in which players, who have other-regarding preferences that depend on the information treatment, learn how to play the game over time. The dependence of preferences on who and what the players observe after each round helps to explain heterogeneity in behavior, and also captures situations in which players switch between very different behavioral strategies across different games. In addition, from an experimental design perspective, these results imply that experimenters can partially control subjects' preferences, even in settings where subjects bring their own, unobservable other-regarding preferences into the lab.

After developing a modified other-regarding preferences model, we use a set of simple criteria to identify and classify each player in a match as a free-rider, an altruist, or a conditional cooperator, based on their choices in the experiments. For the various configurations that emerge in the experiments, players' contributions tend not to converge to the benchmark Nash equilibrium of the stage game by the end of the repeated game. However, the data is consistent with the hypothesis that players learn how to play the game over time. For the circle network structure, we find that players learn differentially, depending on the classification of the other players in their group, and on their relative positions within the network. The main features of the data are closely replicated using learning simulations for each realized configuration of free-riders, altruists, and conditional cooperators. Following the approach in Arifovic & Ledyard (2009), we generate the simulations by merging the Individual Evolutionary Learning model with my modified model of other-regarding preferences.

The results reported in this paper provide some important insights into the way contributions depend on what information individuals observe about others and how the interaction between players is structured. The simple, reactive dynamics of the IEL model, coupled with a model of other-regarding preferences that is keyed to the information players observe between rounds, highlights that learning is an important channel through which information and interaction structure affect contributions. Both of these features of an environment warrant further study, since they may have important implications for organizational or institutional design in settings where individual and group incentives conflict.

There are a number of extensions that follow naturally from the analysis in this paper. First and foremost, the analysis can be replicated for a range of different network architectures and neighborhood structures, in order to test whether the overall network structure affects behavior. While the circle network provides a clearly defined and symmetric neighborhood structure, other network architectures may provoke a variety of different and interesting results. Alternatively, we may find that the effect of players' relative positions on learning is consistent across a range of different networks and neighborhood structures. Further work might also investigate whether the results are sensitive to the marginal return on contributions, and the number of players, both of which are variables that are widely acknowledged to affect contributions.

Another potential extension is to investigate conditional cooperation from the perspective of the reciprocity literature, in which cooperation responds not only to the decisions of others, but also to players' beliefs about others' contributions. In this respect, future research might pursue a model that incorporates some of the insights and techniques used by Fischbacher & Gachter (2011), who actually elicit subjects' beliefs about how others will play in each round of their experiments. Given the uncertainty subjects have about the others participating in the experiment, as well as the considerable heterogeneity among individuals' cooperative preferences, a formal analysis of beliefs may complement the simpler dynamics underlying the IEL model which are used in this paper.

Chapter 4

Learning in Network Public Goods Experiments

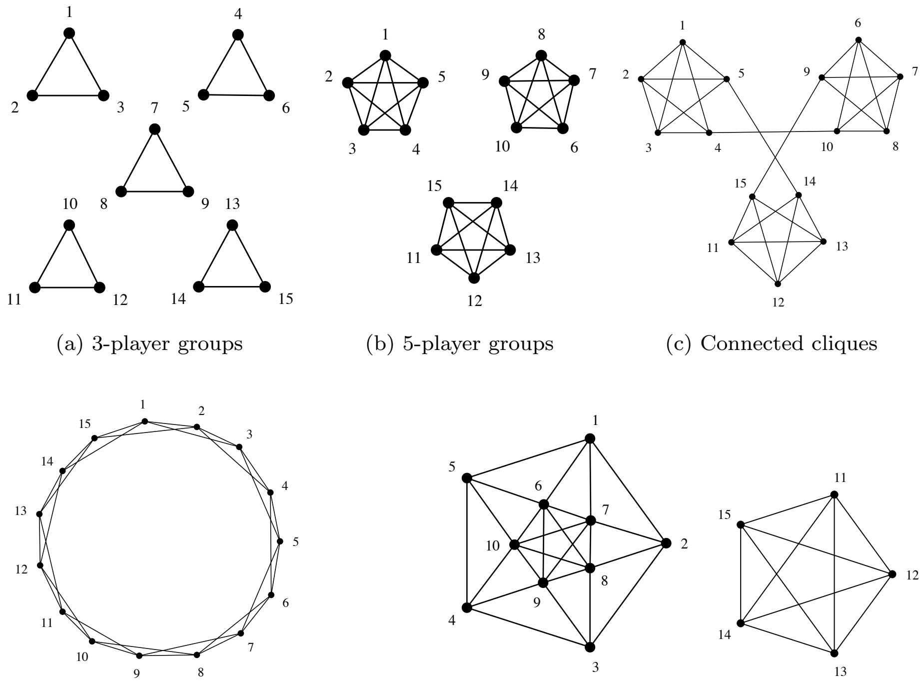
4.1 Introduction

In this chapter, we investigate further the relevance of network structure for contributions in linear public goods experiments. We use the same approach as in Chapter 3 to develop a behavioral model that combines other-regarding preferences with the theory of learning. The approach builds on recent work by Arifovic & Ledyard (2012).[1] The previous chapter focuses on the circle network and the variation in experimental information treatments. In contrast, we focus in this chapter on a single information treatment, played on a set of different network structures. In all of the network environments, players participate in a 10-period repeated network public goods game. After each round, subjects observe the average payoff in their neighborhood, which is defined by network. As in the previous chapter, our behavioral model provides a set of cutoff conditions on a player's preference parameters, which allow us to classify players as free-riders, conditional cooperators, or pure altruists in any given environment.

We study five different network structures in this chapter; two cases in which the players are divided into a collection of complete network components (or cliques), which we call *3-player groups* and *5-player groups*; and three connected, but incomplete networks, called *connected cliques*, *two-step circle*, and *core-periphery* (which is coupled with a separate 5-player clique to keep the total

[1] We refer the reader to the discussion of relevant literature in Chapter 3.

number of players used constant across networks). The networks are shown in Figure 4.1.



(a) 3-player groups (b) 5-player groups (c) Connected cliques

(d) Two-step circle (e) Core-periphery + clique

Figure 4.1: Network Structures

The networks possess several different features. When players are divided into 3-player groups or into 5-player groups, the influences from free-riders and pure altruists are isolated and contained within the player's group. All players in a given clique share a common neighborhood. In contrast, the connected cliques network captures an environment in which all players' decisions have the indirect capacity to affect all other players' decisions, through their neighbors, neighbors' neighbors, and so on. Within each (partial) clique, three players share a common neighborhood, while the other two have distinct neighborhoods that overlap with the other (partial) cliques. In the two-step circle network, each player has a distinct neighborhood. Players' neighborhoods overlap, but with more channels for behavior in one part of the network to indirectly affect behavior in a more distant neighborhood. Finally, in the core-periphery network, players in the periphery have four neighbors, while players in the core have six neighbors, including every other member of the core. This is the only network we consider where some players have more neighbors than others.

This analysis in this chapter is broken into three parts. First, we study the complete network setting with players divided into groups of a certain size. We examine the hypothesis that average group contributions are the same in the 3-player and 5-player environments, keeping the total return to contributions fixed. We find that group contributions are higher in the 5-player groups. However, the average public good consumption for each player is slightly higher in 3-player groups, due to the higher marginal return to contributions. Nevertheless, we emphasize caution in interpreting the findings.

Part of the reason for this caution is that, according to our behavioral theory, changing the size of the groups also changes the incentives for players to free-ride or to give unconditionally. Each environment induces a (potentially) different pair of cutoff conditions in the type space that determine whether an individual has a dominant strategy to free-ride or to contribute his entire endowment. We classify players in each group by applying a set of simple criteria to their decisions in the experiment. While our theory suggests certain restrictions on the classifications across networks, a within-subject analysis of player classification is not consistent with these restrictions. For instance, we can show that a player who is a free-rider in a 3-player group, should also be a free-rider in a 5-player group. Unfortunately, in the data, we consistently observe classifications that violate the predictions of the theory. Although this may be exacerbated somewhat by the small sample size and the coarseness of the classification criteria, it casts a shadow on the predictions of the other-regarding preferences component of our model, at least with respect to the static version of the game.

We also compare average contributions across 5-player groups, connected cliques, and two-step circle networks. In each of these three network environments, every player has exactly four direct neighbors, and the marginal per capita return to contributions is 0.4. As a result, the cutoffs for determining free-riders and altruists are exactly the same in all three networks. In theory then, any differences in average contributions between the networks may be driven by differences in the network structure. We find that average contributions are highest in the 5-player groups, followed by the connected cliques, and lowest in the two-step circle networks. However, again we exercise caution in interpreting the data. A within-subject comparison of classifications across networks is more consistent with the predictions of the theory, although there are still several violations, where a subject is classified differently in the different network environments. While this does not bode particularly well for our theory of other-regarding preferences with respect to the static version of the game, the combined theory of other-regarding preferences and learning, which we turn to next, is more successful.

The second part of our analysis concerns average contributions in groups that consist entirely of conditional cooperators. In these groups, our theory of other-regarding preferences (absent the learning component) allows for a continuum of Nash equilibria, in which all players contribute the same amount. Without an a priori rationale for selecting any one of these Nash equilibria, a natural hypothesis may be that in the repeated setting, players converge to an average contribution equal to half of the endowment. We examine this hypothesis using data from the groups in the experiments with only conditional cooperators and then compare the results with average contributions using simulations from our full behavioral model. In both the data and the learning-based simulations, we find that contributions in these groups decline over time, and converge to levels below half of the endowment. We offer a conjecture for future research, that average contributions are driven down by the learning dynamics. If this is true, then we may be able to make an argument for using the learning dynamics as a method of equilibrium selection.

The third, and most important part of the chapter studies the relationship between learning and the structure of the network. Specifically, we study the hypothesis that learning is contagious throughout the network. This hypothesis follows on from the results obtained for the circle network in Chapter 3, which suggest that free-riders and altruists provide an anchor for learning which spreads first to their neighbors, then on to their neighbors' neighbors, and so on throughout the network. We extend the analysis to the three connected networks; connected cliques, two-step circles, and core-periphery networks. We confirm the result obtained for the circle network, that learning is contagious, especially in the connected cliques and two-step circle networks.

4.2 Experimental Design

We ran six experimental sessions in May and June 2012. All sessions were conducted at the Social Sciences Experimental Laboratory (SSEL) at Caltech using subjects recruited from a pool of undergraduate and graduate students across a range of disciplines. In five of the sessions, we recruited 15 subjects, and in the other session (S3) we recruited 10 subjects.

During the experimental sessions, subjects participated in a set of *matches*, each of which constituted a repeated network public goods game. We used different networks for different matches, but the network structure remained fixed throughout a given match. At the start of each match, the subjects were randomly assigned to a position on the network, so that their positions and their neighbors were fixed for all rounds of the match. Then the subjects played 10 rounds of the same network public goods game. Earnings were denominated in tokens throughout the experimental session, then exchanged to US dollars at the end of the session. Average payouts were approximately US$25, and each session lasted about one hour.

The Game. At the start of each round, we gave subjects 10 tokens and asked them to choose how much of this endowment to contribute towards a project and how much to keep for themselves. Subjects were allowed to contribute any number up to two decimal places, between 0 and 10 (inclusive). A subject's payoff for the round was calculated by adding the tokens they kept for themselves to their earnings from the project. Each subject's earnings from the project were determined by adding together the total number of tokens contributed by the subject and his neighbors in the network, then multiplying this total by a return factor A. Thus, player i's payoff for the round was given by

$$\text{Payoff}_i = 10 - \ i\text{'s contribution} + A \times \left(\sum \ i\text{'s contribution} + i\text{'s neighbors' contributions} \right)$$

After each round, we displayed a summary screen reporting the subject's own contribution, the total contributions made by the subject and his neighbors in the network, the subject's own payoff, and the average round payoff earned by the subject and his neighbors in the network. Thus, information was restricted to the neighborhood level in all matches.

The Networks. In the experiments, we used the five different network structures shown in Figure 4.1. The first two networks are complete networks, which correspond to the standard linear public goods game. The third network, referred to as *connected cliques*, captures an environment in which different clustered groups are connected to each other. The fourth network is a *two-step circle*, in which each player is connected with two neighbors on each side of him, and the fifth network is a *core-periphery* network, which separates the players into a central, clustered group and an outer group of players with fewer connections. The organization of sessions is summarized in Table 4.1.

Table 4.1: Organization and Order of Matches in the Experimental Sessions

In Session 3, the 10 subjects were divided into two 5-player groups, then two 5-player circles, a *paired cliques* network, and a 10-player two-step circle.

	Session					
Network	S1	S2	S3	S4	S5	S6
3-player groups, $A = \frac{2}{3}$	1	1	–	–	–	–
5-player groups, $A = 0.4$	2	2	1* & 2*	1	1	1
Connected Cliques, $A = 0.4$	3	3	3*	3	3 & 6	3 & 6
Two-Step Circle, $A = 0.4$	4	4	4*	2	2 & 5	4 & 7
Core-Periphery + Clique, $A = 0.4$	–	–	–	–	4 & 7	2 & 5
Total Matches	4	4	4*	3	7	7

4.3 The Model

We use a model which is based on the approach taken by Arifovic & Ledyard (2012) for linear voluntary contributions games without networks. They merge a standard, outcome-based other-regarding preferences model with a theory of learning. We extend the model to allow for the network structure of the environment.

The agents have other-regarding preferences over outcomes. We provide a utility representation for these preferences that consists of three components. First, the agents derive utility from their own payoff. Second, they exhibit a social preference that depends on the average payoff earned by their neighbors in the network. Third, they have a preference for fairness to self, in the sense that they do not like to be earning less than the average payoff in their neighborhood. This third component may be described as envy, or as one-sided fairness.

Although we assume a common utility functional form, we allow for heterogeneity among the agents through two parameters that represent an agent's intensity of social preference, and his intensity of preference for fairness. The equilibrium behavior induced by this utility function varies between complete free-riding, unconditional full contribution, and conditional cooperation. An important point of distinction from the earlier literature is that these strategies do not emerge solely from the characteristics of the agents, but rather from the confluence of their types with the parameters of the environment.

The theory of other-regarding preferences generates static equilibrium predictions, but provides little insight into the dynamics of behavior in repeated public goods experiments. Thus, we combine the model with a theory of learning, which argues that agents are boundedly rational, and that they learn how best to play over the course of the repeated game. Another popular approach argues that the dynamics are driven by strategic behavior, rather than learning. However, this alternative approach rests on strong assumptions of common knowledge about both beliefs and rationality, which are difficult to justify in 10-period repeated games with many players. Thus, although there is some evidence consistent with strategic play in the last few rounds of repeated public goods experiments, we focus on the learning approach.

4.3.1 Game Notation

Let $I = \{1, ..., n\}$ be the set of players in the population, each with an endowment $\omega_i = 1$. Each player must choose how much of his endowment to consume privately, and how much to allocate towards the public good. We denote player i's contribution to the public good by $y_i \in [0, 1]$ and the profile of all players' contributions by $\mathbf{y} = (y_i)_{i \in I}$.

The players are connected through a network, denoted by an $n \times n$ matrix, G. An entry $G_{ij} = 1$ indicates that player i is connected to player j. We restrict attention to undirected networks, so that $G_{ij} = G_{ji}$ for all pairs i, j, and adopt the convention $G_{ii} = 0$. Let $N_i(G)$ be the set of neighbors for player i in the network G. Formally,

$$N_i(G) = \{j \in I | G_{ij} = 1\}. \tag{4.1}$$

Let $k_i(G) = |N_i(G)|$ be player i's degree (or number of neighbors). The level of public good enjoyed by player i depends on his own contribution, the contributions made by his neighbors, and his marginal per capita return to contributions, $\frac{M}{k_i+1}$. That is,

$$Y^i(G, \mathbf{y}) = \frac{M}{k_i + 1} \cdot \left[y_i + \sum_{j \in N_i(G)} y_j \right]. \tag{4.2}$$

Player i's payoff is given by

$$\pi_i(G, \mathbf{y}) = 1 - y_i + Y^i(G, \mathbf{y}). \tag{4.3}$$

For each network, we assume that M is such that $1 < M < k_i + 1$ for each $i \in I$. This ensures that the self-interested players face the familiar social dilemma, in which social and individual incentives conflict.

4.3.2 Preferences

Players have other-regarding preferences over outcomes, which we capture using the utility function

$$U_i(G, \mathbf{y}) = \pi_i(G, \mathbf{y}) + \beta^i \cdot \sum_{j \in N_i(G) \cup \{i\}} \frac{\pi_j(G, \mathbf{y})}{k_i + 1} - \gamma^i \cdot \max \left\{ 0 \, , \sum_{j \in N_i(G) \cup \{i\}} \frac{\pi_j(G, \mathbf{y})}{k_i + 1} - \pi_i(G, \mathbf{y}) \right\}. \tag{4.4}$$

The two parameters β^i and γ^i represent, respectively, player i's relative concern for his neighborhood and his relative concern with the fairness of his own payoff compared to social payoffs. We assume that $\beta^i \geq 0$ and $\gamma^i \geq 0$. There is a very natural interpretation of the three components to this utility function. The first term represents personal, or self-interested preference through the player's own payoff. The second term corresponds to an altruistic or social preference that extends over the player's neighborhood in the network, and the third term reflects the degree of envy experienced by the player, when his own payoff is less than the average payoff in his neighborhood.

4.3.3 Learning

The agents in our model are boundedly rational individuals, who learn by reacting in a 'best response' manner to outcomes in past periods of play. We adopt the Individual Evolutionary Learning (IEL) model developed by Arifovic & Ledyard (2012). IEL is particularly well-suited to repeated game environments with multiple rounds, multiple players, and continuous action spaces. Furthermore, it is a relatively simple model, robust to the specified value of model parameters, and specifies individual-based, reactive dynamics, in contrast with theories of social learning.

Individual Evolutionary Learning

The IEL model consists of several components. In any round, each agent retains a finite set of remembered actions. Each agent selects an action from his remembered set according to a probability measure that evolves with the set. After each round, the set of remembered actions is updated through experimentation and replication. The process is initialized by randomly populating the set of remembered actions and assuming that each element is chosen with equal probability. The experimentation procedure allows agents to make mistakes and to consider actions that might not otherwise be discovered. The replication and selection processes are designed to favor actions that would have generated higher payoffs, according to an agent-specific *forgone utility function.* The forgone utility calculations provide the channel through which we merge the other-regarding preferences with the learning dynamics.

Experimentation involves replacing (with some small probability) some of the elements of the remembered set with an action chosen at random from the entire action space. *Replication* makes, for each slot in the remembered set, a comparison between two remembered alternatives and replaces the action in the slot with the alternative that has a higher forgone utility for the previous round. *Selection* chooses an action from the remembered set proportionately to the forgone utility it would have earned in the previous round.

Learning Simulations. For all simulations in this chapter, we used a remembered set of size $J = 50$, and a probability of experimentation $\rho = 0.033$. New alternatives were selected according to a normal distribution, conditional on the set of possible contributions $[0, 10]$, with mean equal to the old alternative and a standard deviation equal to $\sigma = 5$. We ran 500 iterations for each simulation, and a simulation for each individual match in the dataset, based on the classified configuration of players. Only the contributions of the conditional cooperators were simulated, while we set classified free-rider and altruist contributions to be 0 tokens and 10 tokens, respectively, in each round. For the conditional cooperators, each iteration of a simulation drew a new pair of preference parameters from a range consistent with conditional cooperation. We incorporated the different network structures into separate simulation code, although the underlying parameter values were the same across all networks.

4.3.4 Equilibrium Behavior

It will be useful to derive the static equilibrium predictions for the one-shot game as a benchmark for individual behavior. Fix the network structure G, and the set of types $(\beta^i, \gamma^i)_{i \in I}$. The equilibrium strategies are derived from the players' marginal utility functions $\frac{\partial U_i}{\partial y_i}$.

Lemma 4.3.1. *Player i's marginal utility is always negative if*

$$\beta^i < \frac{(k_i + 1)(1 - A_i)}{\sum\limits_{j \in N_i(G)} A_j - (1 - A_i)}. \tag{4.5}$$

Lemma 4.3.2. *Player i's marginal utility is always positive if*

$$\beta^i > \frac{(k_i + 1)(1 - A_i)}{\sum\limits_{j \in N_i(G)} A_j - (1 - A_i)} + \gamma^i \left(\frac{\sum\limits_{j \in N_i(G)} A_j + k_i(1 - A_i)}{\sum\limits_{j \in N_i(G)} A_j - (1 - A_i)} \right). \tag{4.6}$$

These two lemmas have a simple interpretation. If a player's type (β^i, γ^i) satisfies the inequality 4.5 in Lemma 4.3.1, his dominant strategy is to contribute zero. In other words, a player with sufficiently low social preference will always free-ride. On other hand, if (β^i, γ^i) satisfy the inequality 4.6 in Lemma 4.3.2, player i's dominant strategy is to contribute his entire endowment. That is, a player with sufficiently high social preference relative to his preference for fairness, will always contribute everything towards the public good.

If player i's type parameters do not satisfy either of conditions 4.5 or 4.6, then utility is maximized by choosing a contribution level that minimizes the inequality between i's payoff and his neighborhood average payoff. For any game, the equilibrium strategies for *free-riders* and *pure altruists* are determined by condition 4.5 and condition 4.6, respectively. Then the equilibrium strategies for *conditional cooperators* are derived by solving a system of linear inequalities, subject to the boundary constraints on individual contributions. Whenever there exists at least one free-rider or altruist in the group, the Nash equilibrium is unique. However, in the special case when every player in the network is a conditional cooperator, there are a continuum of equilibria, in which each player contributes the same amount $x \in [0, \omega^i]$.

4.4 Hypotheses

We investigate three natural hypotheses in this paper. First, we study group contributions in 3-player groups and 5-player groups, when the return to total contributions is kept fixed. We are interested in whether the lower marginal per capita return to contributions balances out the higher number of players in the larger group thereby inducing the same average group contributions. Second, we look at behavior in groups that are composed only of conditional cooperators. In theory, there are a continuum of equilibria in these environments, where each player contributes the same amount. In expectation, it seems natural to argue that average contributions in these groups converge to half of the endowment. Finally, we investigate whether learning is contagious throughout the networks. Our hypothesis of learning contagion builds on the results obtained for the circle network in Chapter 3. We outline these hypotheses below.

Hypothesis 1. *Average group contributions are the same in 3-player groups as in 5-player groups.*

The intuition for this hypothesis is as follows. In both environments, the total return to group contributions is equal to 2. The marginal return to contributions is $\frac{2}{3}$ in the 3-player group, and 0.4 in the 5-player group. Thus, we might expect higher average individual contributions in the 3-player group, but there are fewer players contributing to the group total. A natural expectation may be that total group contributions are, on average, the same in both environments.

However, there are also reasons to doubt this hypothesis. First, according to our model of other-regarding preferences, the cutoffs in the type space that determine who is a free-rider, who is a conditional cooperator, and who is a pure altruist, are different for the two different group sizes. It is relatively straightforward to show that any player with an incentive to free-ride in the 3-player environment also has an incentive to free-ride in the 5-player environment. From Lemma 4.3.1, if $\beta^i < \frac{3(1-\frac{2}{3})}{3(\frac{2}{3})-1} = 1$ in the 3-player group, player i is a free-rider. In the 5-player group, the cutoff condition is $\beta^i < \frac{5(1-0.4)}{5(0.4)-1} = 3$, which is higher than in the 3-player group. On the other hand, anyone with a dominant strategy to contribute everything in the 5-player group also has a dominant strategy to contribute everything in the 3-player group. As a result, it is possible that the same distribution of types (underlying preference parameters) can lead to a higher percentage of free-riders and lower percentage of pure altruists in the 5-player groups.

Second, even if the distribution of free-riders, pure altruists, and conditional cooperators is the same in the two networks, the configuration of types on the network may have a significant impact on contributions. For example, suppose that there are 5 subjects who are free-riders in both the 3-player and 5-player environments. Suppose further that there is exactly one of the five free-riders in each of the five 3-player groups. Then the Nash equilibrium for each group is zero contributions by all players. Then suppose that all five free-riders are collected in one of the three 5-player groups, while the other two consist of all conditional cooperators. In this case, the Nash equilibrium for the group of free-riders is zero contributions by everyone, but in the other two groups, there are a continuum of equilibria in which every player contributes the same amount. As a result, since all the free-riders are confined within one of the 5-player groups, the average contribution may be higher in the 5-player environment than in the 3-player environment. There exists an analogous example for which the reverse is true.

Finally, we argue that if players are learning how to play the game over time, the learning dynamics may depend on the number of players in the group, or more generally on the structure of the network. For instance, in 3-player groups, there may be less feedback from conditional cooperators who are attempting to coordinate with each other on the utility-maximizing Nash equilibrium contribution.

Hypothesis 2. *In groups that consist only of conditional cooperators, average contributions converge to half of the endowment.*

In any group consisting of all conditional cooperators, the other-regarding preferences model does not provide very helpful predictions. In these cases, there are a continuum of Nash equilibria, in which all players contribute the same amount, $x_i = x \in [0, \omega], \forall i \in I$. This complicates the coordination problem for the conditional cooperators, since they may each attempt to coordinate on a different equilibrium contribution. This also makes it difficult for us as analysts to determine average contributions in all-conditional cooperator groups.

A natural first-guess is that, across groups with only Cs, the average contribution is half the endowment. This simply reflects an expectation over the action space, combined with the following prediction. Higher-contributing players respond to the lower contributions by reducing their contributions over the course of the repeated game, while the lower-contributing players respond by increasing their contribution. We analyze contributions in the experiment for groups with all conditional cooperators, both overall and by network, then compare these against the hypothesis that they converge, on average, to half of the endowment.

Hypothesis 3. *Learning is contagious throughout the network.*

This hypothesis builds on the results obtained for the circle network in Chapter 3 of this dissertation. There, we found that players exhibit differential and contagious learning across the network, depending on their position relative to any free-riders or altruists in the group. The intuition for contagious learning rests on the idea that free-riders and altruists provide an anchor for learning, since their decisions are stable from one round to the next. Conditional cooperators who are directly connected to a free-rider or altruist are able to refine their strategies using this stability, while those who are only connected to other conditional cooperators are faced with a difficult coordination problem. However, once some conditional cooperators begin to learn how best to play the game, the stability provided by the free-riders or altruists in the group spreads to other conditional cooperators, along the links in the network. This is an important result, since it suggests that network structure may have a significant effect on actual contributions, through its effect on learning by conditional cooperators. In this paper, we further investigate the hypothesis that learning is contagious, for a different set of networks.

4.5 Experimental Results

4.5.1 Player Classification

In this section, we summarize the experimental results. First, we classify the players in the experiments. We classify the players in each match, using a simple set of criteria. Each player is classified as a *free-rider* (F), *pure altruist* (A), or *conditional cooperator* (C), based on their decisions in the match. We apply the criteria listed below to decisions made in rounds 3 through 10. A player receives classification

- F if he contributes 0 in at least seven rounds;
- A if he contributes 10 in at least seven rounds;
- C otherwise.

In Table 4.2, we summarize the classifications for players in each of the network environments. Figure 4.2 shows the proportion of the three classifications by network.

Table 4.2: Classifications by Network

Network	Classification F	C	A	Total
3-player groups	5	19	6	30
5-player groups	13	81	11	105
Connected Cliques	16	79	10	105
Two-step Circle	22	78	5	105
Core-Periphery	1	29	10	40

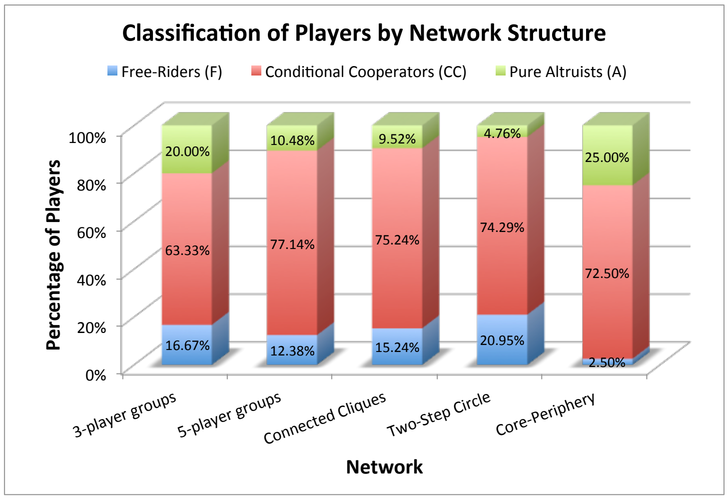


Figure 4.2: Proportion of players in each classification for the different network environments

4.5.2 Group Contributions in Complete Networks

Next, we consider Hypothesis 1. In Figure 4.3, we plot the average group contributions for 3-player and 5-player groups. The average for the 5-player groups is higher in all rounds. However, the average level of public good consumption is very similar in the two environments, since the rate of return is lower in the 5-player groups. This similarity is illustrated in Figure 4.4.

These findings should be interpreted with care, since there may be a different distribution of free-riders, pure altruists, and conditional cooperators in the two environments. In the 3-player groups, a player with altruistic parameter $\beta^i < 1$ has a dominant strategy incentive to free-ride, while in the 5-player groups, any player with $\beta^i < 3$ has the same incentive to free-ride. Thus, if underlying types (utility parameters) are stable across environments, a player who free-rides in a 3-player group also free-rides in a 5-player group. Similarly, any pure altruist in a 5-player group will also be a pure altruist in a 3-player group. These differences in the free-rider and pure altruist

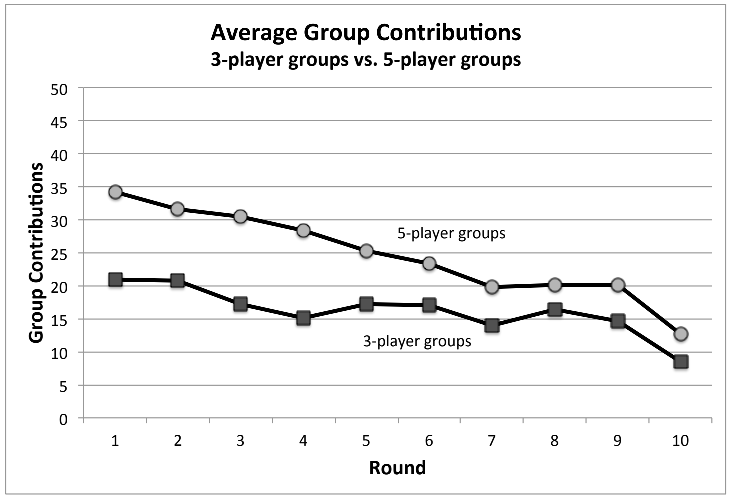


Figure 4.3: Average group contributions in 3-player and 5-player groups

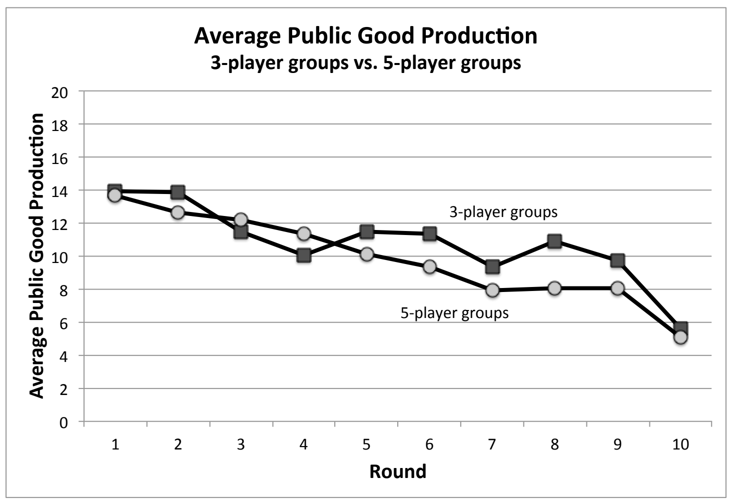
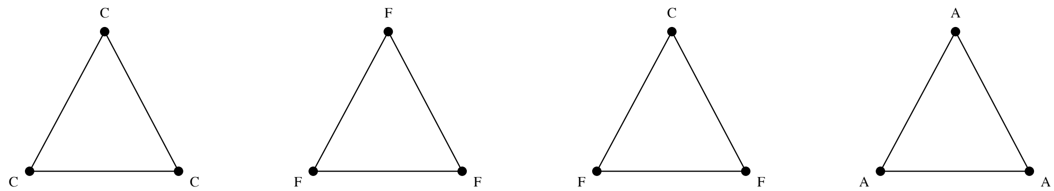


Figure 4.4: Average public good consumption in 3-player and 5-player groups

cutoffs could, in theory, induce lower average individual contributions in 5-player groups than in 3-player groups.

Although the sample sizes are different, we find a higher proportion of free-riders in the 3-player groups than in the 5-player groups. This is not consistent with our theory, although it is possible that the reversal is driven by differences in the sample sizes. Nevertheless, by analyzing the within-subject classifications for 3-player and 5-player groups, we can determine whether players are stable across environments. Unfortunately, the within-subject analysis is also fairly inconsistent with the predictions of the model with respect to cutoffs in the type space.



(a) **C C C (6 groups):** There are a continuum of equilibria in which all players choose the same contribution.

(b) **F F F (1 group):** In equilibrium, all players contribute 0.

(c) **F C F (1 group):** In equilibrium, all players contribute 0.

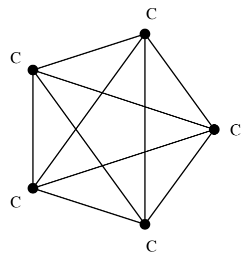
(d) **A A A (2 groups):** In equilibrium, all players contribute 10.

Figure 4.5: 3-player group configurations and their Nash equilibria

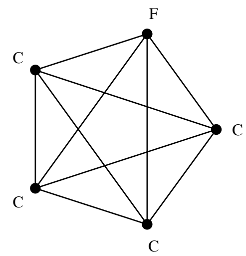
Of the five subjects who are classified free-riders in their 3-player group, only one is also classified as a free-rider in his 5-player group. The other four are each classified as conditional cooperators, which is inconsistent with the order of the cutoffs, derived from our theory of other-regarding preferences. The contributions of those conditional cooperators are, in a few cases, only marginally excluded from a free-rider classification, but in general, this suggests that either the theory does not predict well, or that subjects are not particularly stable across environments. Similarly, the only subject who is classified as a pure altruist in the 5-player group, and who also played in a 3-player group, is classified as a conditional cooperator in the latter, contrary to theoretical predictions.

In addition to these concerns, group contributions can also depend on the configuration of types in the group. For instance, even if within-subject classifications were perfectly robust and consistent with the theory, the way the players are configured among the groups may change the comparative average contributions. On examining the configurations in the two environments, we find that the set of configurations is quite similar for 3-player groups as for 5-player groups. The configurations, and the corresponding Nash equilibria, are illustrated in Figure 4.5 (for 3-player groups), and Figure 4.6 (for 5-player groups).

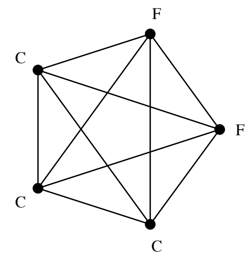
In the 3-player groups, there are two AAA configurations, one FFF configuration, and one FCF configuration. On average, contributions from these four matches should be half of the endowment. The other six groups consist entirely of conditional cooperators. In these configurations, there are a continuum of equilibria, in which every player contributes the same amount. Among the 5-player groups, there are similarly balanced configurations, with six groups composed of only Fs and Cs,



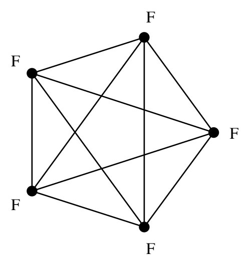
(a) **C C C C C (8 groups):** There are a continuum of equilibria in which all players choose the same contribution.



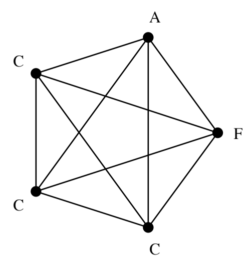
(b) **F C C C C (4 groups):** In equilibrium, all players contribute 0.



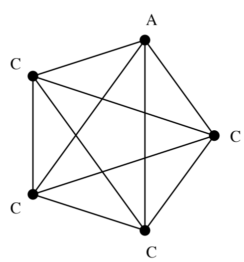
(c) **F F C C C (1 group):** In equilibrium, all players contribute 0.



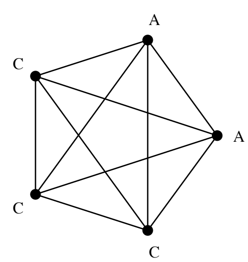
(d) **F F F F F (1 group):** In equilibrium, all players contribute 0.



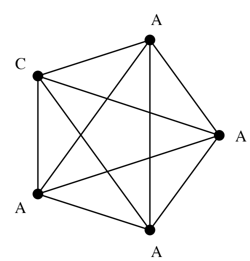
(e) **F A C C C (2 groups):** In equilibrium, the Cs all contribute 5, the F contributes 0, and the A contributes 10.



(f) **A C C C C (3 groups):** In equilibrium, all players contribute 10.



(g) **A A C C C (1 group):** In equilibrium, all players contribute 10.



(h) **A A A A C (1 group):** In equilibrium, all players contribute 10.

Figure 4.6: 5-player group configurations and their Nash equilibria

five groups composed of only As and Cs, two groups with one F, one A, and three Cs, and eight groups that consist of all Cs. The groups with free-riders (and Cs) and the groups with altruists (and Cs) should average each other out to half the endowment.

4.5.3 Groups Composed Solely of Conditional Cooperators

As discussed in Section 4.3.4 and Section 4.4, there are a continuum of equilibria in matches composed of all conditional cooperators. A natural expectation is that the average contributions in these matches converge to half of the endowment. However, in our learning simulations, we find that contributions in these matches tend to converge to values significantly below half of the endowment. We conjecture that learning dynamics drive the average contributions down, and that different networks generate different learning dynamics. It may be possible to use these learning dynamics as a method of equilibrium selection, particularly as it explains the well-documented decay in average contributions over time, even for groups that consist solely of conditional cooperators. We leave a deeper analysis of this conjecture for future research.

In our data, there are six groups of all conditional cooperators in 3-player groups, eight groups of all conditional cooperators in 5-player groups, and one group of all conditional cooperators on a two-step circle. We show the average contributions in learning simulations for the 3-player groups and 5-player groups in Figure 4.7. We also plot the average contributions in the data for those groups with all conditional cooperators. For 5-player groups, we find that the data follows the simulations quite closely, as average contributions decline over time to a level below half of the endowment. The simulation averages converge to approximately 2.5 tokens, and while the data average starts at more than half of the endowment, it falls into line with the simulation average by round 5. On the other hand, the data average for 3-player groups begins and remains higher than half the endowment for most rounds, while the simulations decline steadily below 5 tokens and converge to approximately 3.3 tokens by the end of the 10 rounds. Nevertheless, the data average also appears to decline over time. The observation that average contributions are much lower for groups composed of all Cs in 5-player groups than in the 3-player groups may partly explain why we find average group contributions to be slightly lower in 5-player groups than in 3-player groups.

Although there is not enough data to do the same analysis for the other networks, Figure 4.8 shows the predictions of the learning simulations for the connected cliques, two-step circle, and core-periphery networks when all players are conditional cooperators. All three networks exhibit

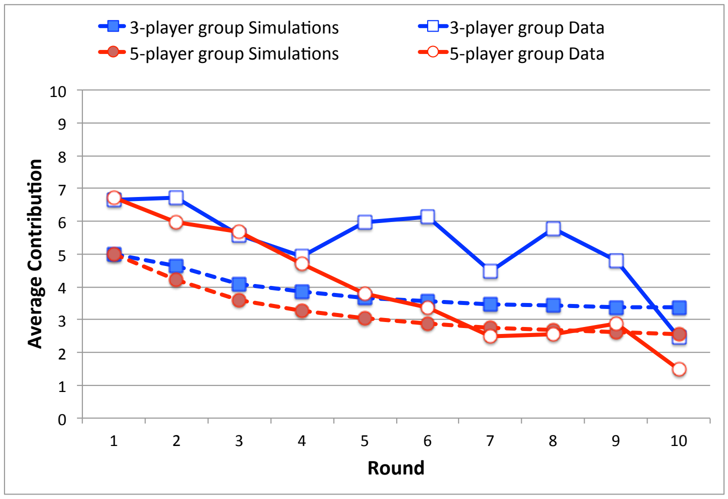


Figure 4.7: Average contributions in groups with all conditional cooperators

virtually identical patterns of average contributions, beginning at 5 tokens and declining to 4 tokens by the end of the simulated game.

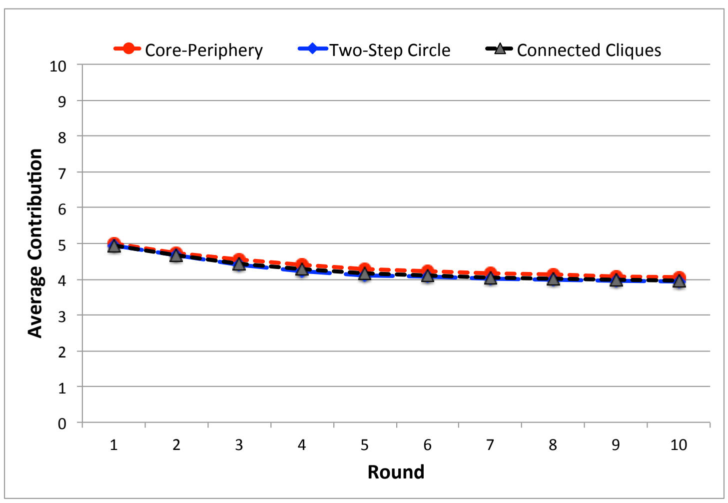


Figure 4.8: Average simulated contributions in groups with all conditional cooperators

4.5.4 Average Contributions in Networks with 5-Player Neighborhoods

In this section, we compare contributions across the three network environments in which every player is part of a 5-player neighborhood, and the marginal per capita return to contributions is 0.4. These networks are the 5-player groups (or cliques), the connected cliques network, and the two-step circle network. The cutoffs in the type space for determining free-riders and altruists are exactly the same in all three networks, so in theory, the distribution of players should be the same. We do allow for some expected differences, since there are some players who play on all three networks twice, while others only play on the connected cliques and two-step circle networks twice.

Figure 4.9 illustrates the average contributions in each of the three networks. The data indicate average contributions are highest in the 5-player groups, followed by the connected cliques, and lowest in the two-step circle networks.

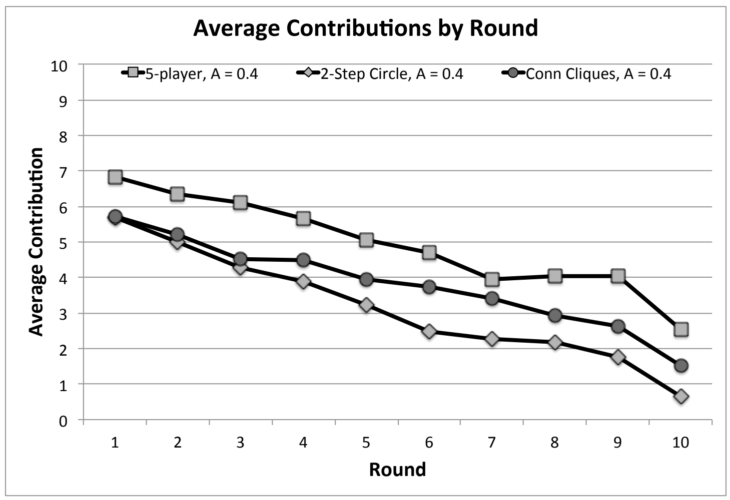


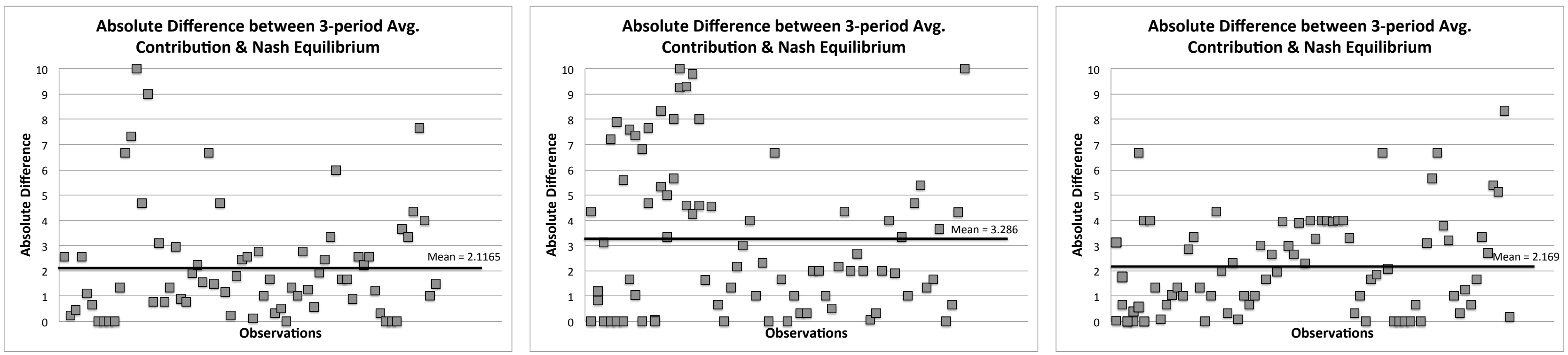
Figure 4.9: Average contributions in 5-player groups, connected cliques, and two-step circle networks

As in the previous subsection, we exercise caution in interpreting this data. First, although the proportion of conditional cooperators is very stable across the three networks, there are progressively fewer altruists and more free-riders as we move from 5-player groups, to connected cliques, to two-step circles. It is difficult to rule out the possibility that these differences in the distribution of types are responsible for the differences in average contributions. As with the comparison of the complete networks in the previous subsection, the configuration of those types on the network may also have important effects on contributions.

A within-subject comparison suggests somewhat more consistency in classifications than we found between 3-player and 5-player groups. Our model predicts that a subject should be classified the same way in all three networks. There are 26 (out of 75) subjects who are classified differently across the three networks. Five of the subjects who participated twice in the connected cliques and two-step circle networks, were classified differently in just one of their 5 matches. Furthermore, in the first session, 7 out of the 9 inconsistent subjects were classified C in both the 5-player groups and the connected cliques, then classified F in the two-step circle environment. This dramatic shift towards free-riding in the two-step circle, which was the final match played by the subjects in that session, raises the possibility that behavior in later matches is influenced by earlier matches. In our design, we varied the order of matches across most sessions, although the setup makes it difficult to prevent all order effects from influencing the data. Although this is an improvement on the consistency of classifications between the 3-player and 5-player groups, it does not provide much support for the predictions of our theory.

4.5.5 Learning in Networks with 5-Player Neighborhoods

In Figure 4.10, we compare the absolute difference between individuals' contributions (averaged over rounds 8–10) and their corresponding Nash equilibrium contribution. For matches composed of all conditional cooperators, we used the absolute difference between the 3-period data average and the average contribution in round 10 of the learning simulations for that network. The average difference is highest for the connected cliques, and almost equal for 5-player groups and two-step circles. This suggests that perhaps conditional cooperators are slower to learn in connected cliques, than in 5-player groups or two-step circles. However, it seems necessary to acknowledge that learning depends crucially on the realized distribution of types and their configuration on the network, rather than the underlying structure of the network itself. Thus, we focus the rest of the analysis on learning at the individual match level.



(a) 5-player groups (b) Connected cliques (c) Two-step circle

Figure 4.10: Absolute differences between 3-period average contribution (rd 8–10) and Nash equilibrium. The mean is shown by the horizontal line at 2.1165 (5-player), 3.286 (connected cliques), and 2.169 (two-step circle).

4.5.6 Learning Contagion in the Incomplete Networks

In this section, we examine the evidence of contagious learning in the connected cliques, two-step circle, and core-periphery networks. We focus on the data at the individual match level, and compare actual contributions with the learning simulations for each realized configuration.

Evidence from Connected Cliques

Session 1, Match 3. In this match, there are four free-riders, located at nodes 2, 6, 7, and 8, as shown in Figure 4.11. In the figure, we use large, red circles to denote the free-riders, then use color and size changes to indicate the relative speeds at which we hypothesize the players are learning. The slower the player learns, the smaller the circle. The color of the node changes from red (free-riders) and orange (fast) to a pale cream, then to progressively darker shades of blue (slowest).

In this configuration, the hypothesis that learning is contagious suggests that players 9 and 10 should converge relatively quickly towards the Nash equilibrium contribution of zero. Players 1, 3, 4, and 5 (player 2's neighbors) should be close behind, followed by players 14 and 15, while players 11, 12, and 13 are most likely the slowest to converge towards zero.

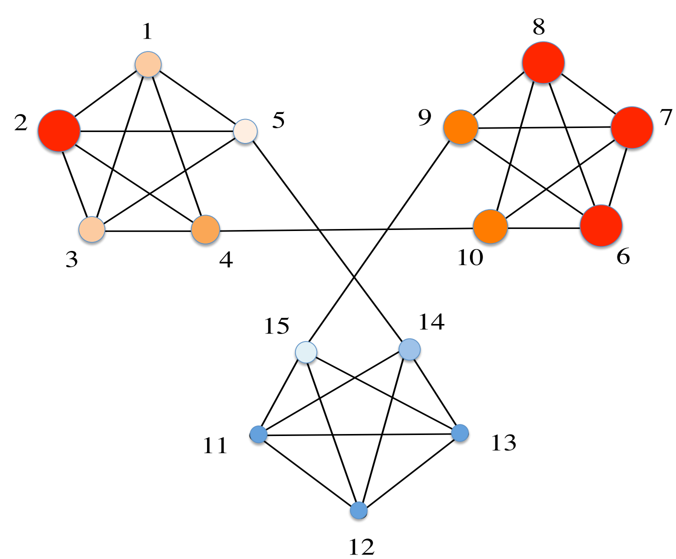
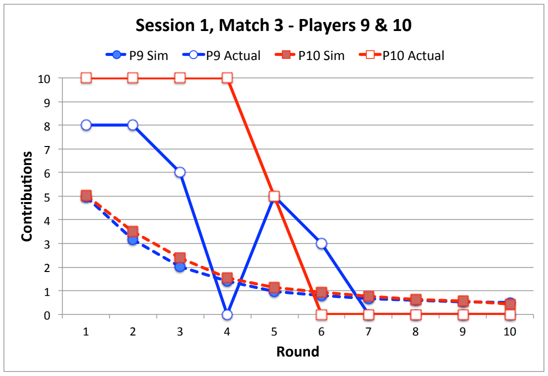
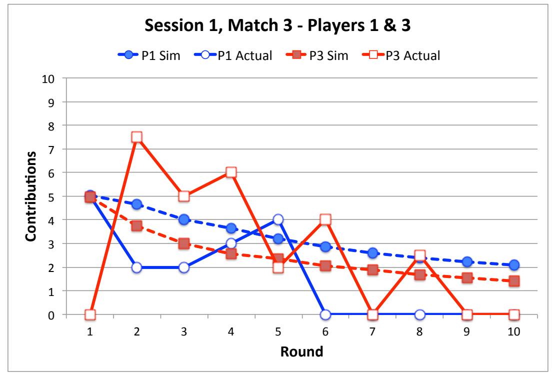


Figure 4.11: Configuration in Session 1, Match 3

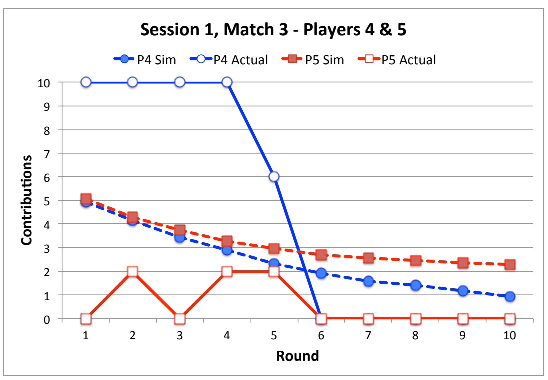
Our intuition is validated by both the learning simulations and the experimental data for this configuration. The free-riders provide an anchor, which facilitates learning by their immediate neighbors. The stabilizing effects of the anchors then spread throughout the rest of the network. We show the individual contributions and the simulated contributions for each conditional cooperator in Figure 4.12.



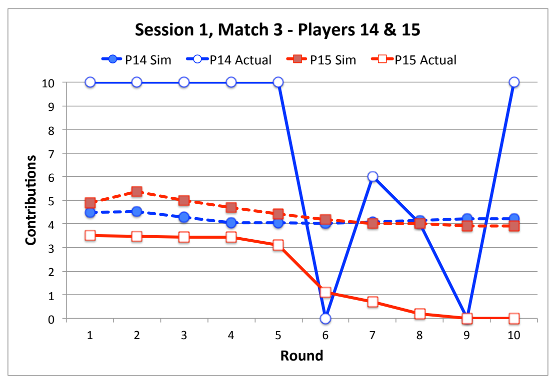
(a) Players 9 & 10



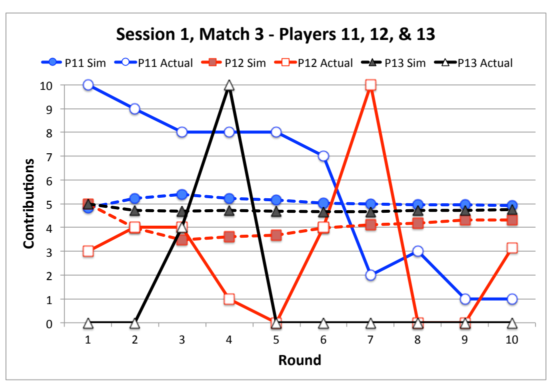
(b) Players 1 & 3



(c) Players 4 & 5



(d) Players 14 & 15



(e) Players 11, 12, & 13

Figure 4.12: Experimental data and learning simulations in Session 1, Match 3

Session 2, Match 3. In this match, there is one altruist (player 5) and one free-rider (player

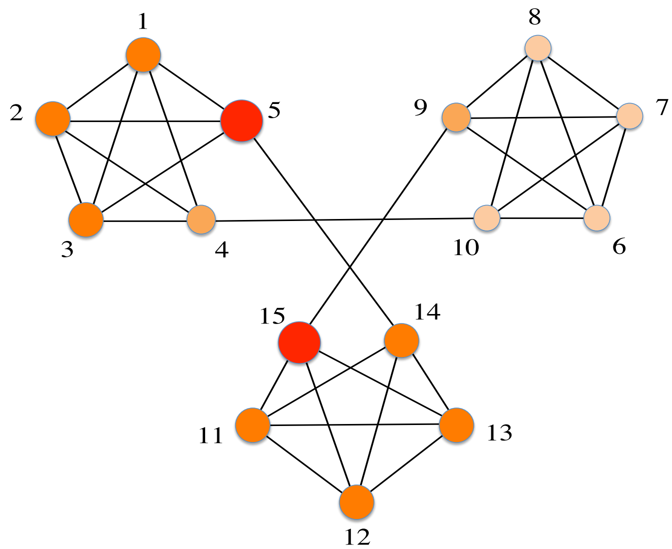


Figure 4.13: Configuration in Session 2, Match 3

15). We solve for the Nash equilibrium from the appropriate set of inequalities, subject to the boundary constraints on contributions, and obtain

$$y_1^* = 10,\ y_2^* = 10,\ y_3^* = 10,\ y_4^* = 9.654,\ y_5^* = 10$$

$$y_6^* = 5.255,\ y_7^* = 5.255,\ y_8^* = 5.255,\ y_9^* = 3.483,\ y_{10}^* = 6.782$$

$$y_{11}^* = 0,\ y_{12}^* = 0,\ y_{13}^* = 0,\ y_{14}^* = 2.5,\ y_{15}^* = 0.$$

There are several symmetric players in this network. Player 1, 2, and 3 are in identical positions as one another. Likewise, players 6, 7, and 8 face the same incentives as each other, and players 11, 12, and 13 are symmetric with one another. In the network, we expect players 1, 2, and 3 to learn quite quickly, since their contributions are well-anchored against the unconditional giving of player 5. Similarly, players 11, 12, and 13 learn from their free-rider neighbor, player 15, but are also reinforced by player 14, who is one of player 5's neighbors. Spreading out from those two partial cliques, we expect player 4 to be slower to learn their Nash equilibrium contribution, while player 10 and players 6, 7, and 8 are the slowest to learn, as they are isolated from the stabilizing effects of a free-rider or altruist. In Figure 4.14, we show the actual and simulated contributions for each player over all 10 rounds.

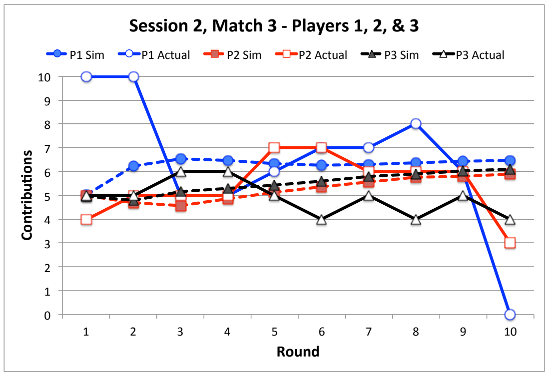
As in Session 1, Match 3, there is some evidence of contagious learning. Players 11, 12, and 13 converge quickly towards their Nash equilibrium contribution (zero) in both simulations and data, while players 6, 7, and 8 are relatively far from their equilibrium contribution in the data. However, player 4 and player 14 behave differently from the simulations. Since players 1, 2, and 3 do not converge to their equilibrium contributions particularly quickly, player 4 is equally slow, particularly since his other neighbor, player 10 is also isolated and contributes less than his equilibrium contribution. On the other hand, player 14, whose equilibrium contribution is 2.5 tokens, converges relatively quickly to a low contribution, even converging to zero in the last two rounds. This reflects the relative speed with which players 11, 12, and 13 learn, which strongly affirms the hypothesis of contagious learning. The evidence from other matches generally provides good support for the hypothesis of contagious learning. We relegate the analysis of these other matches to Appendix ...

Evidence from Two-Step Circles

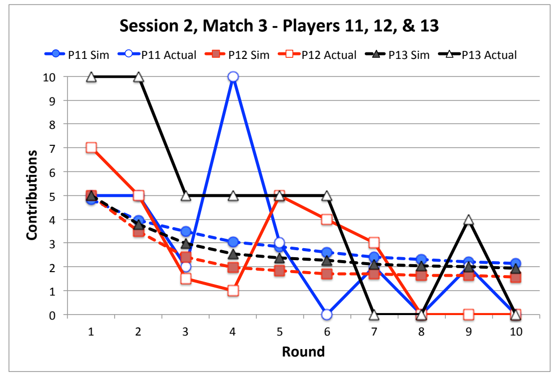
Session 1, Match 4. In this match, there are 11 free-riders, and only 4 conditional cooperators (at nodes 2, 3, 9, and 12). The Nash equilibrium contribution for every player is 0. We expect all four players to converge quite quickly to zero contributions. If anything, players 9 and 12 may converge slightly faster, since they each have four free-rider neighbors, while players 2 and 3 are connected to each other, and three free-riders each. We display the configuration and conjecture some differential learning for the network in Figure 4.15.

Indeed, as is shown in Figure 4.16, all four conditional cooperators converge to the Nash equilibrium contribution (zero) by the last few rounds. Player 12 is marginally quicker to reach and stay at zero contributions, although there is not much difference in the speed of learning, since all of them are well connected to the free-riders in the network.

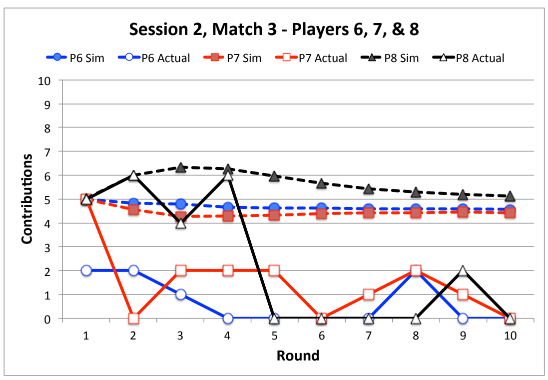
Session 2, Match 4. In this match, there are only 2 free-riders, located at nodes 5 and 6. All other players are conditional cooperators. The Nash equilibrium contribution for every player is 0. Since the free-riders are located right next to each other in the circle, there is a very natural pathway



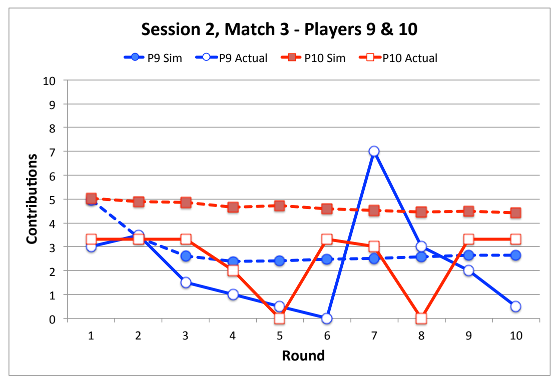
(a) Players 1, 2, & 3



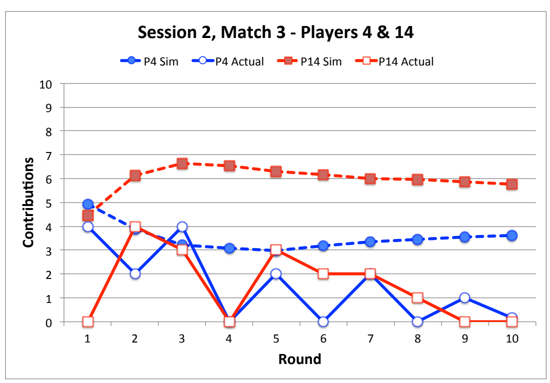
(b) Players 11, 12, & 13



(c) Players 6, 7, & 8



(d) Players 9 & 10



(e) Players 4 & 14

Figure 4.14: Experimental data and learning simulations in Session 2, Match 3

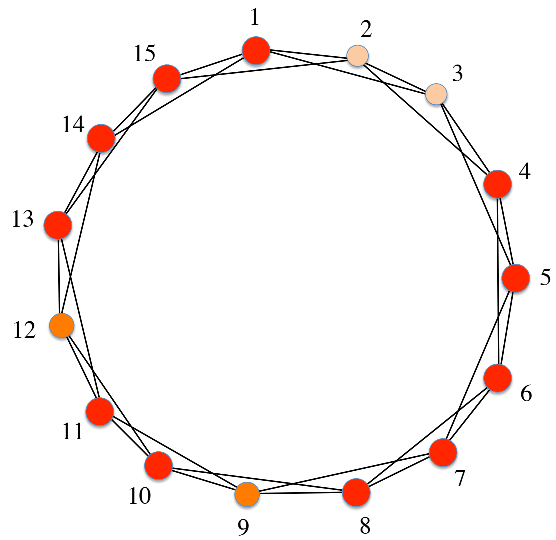
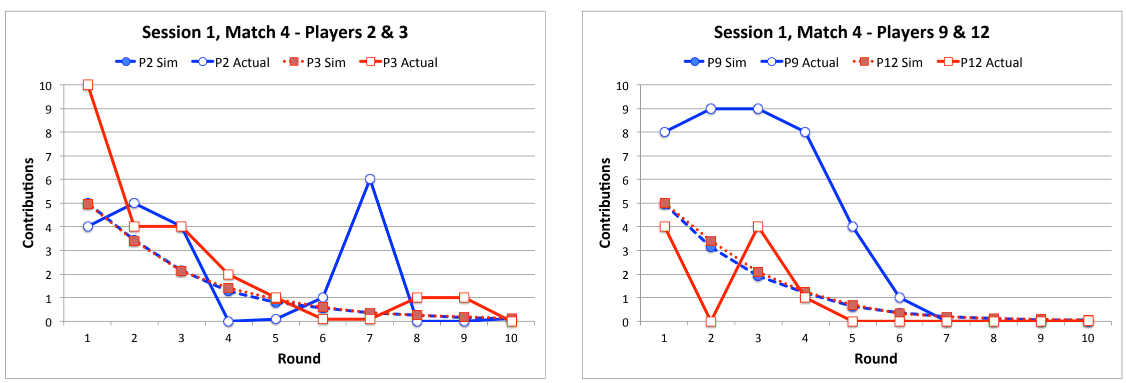


Figure 4.15: Configuration in Session 1, Match 4



(a) Players 2, & 3

(b) Players 9 & 12

Figure 4.16: Experimental data and learning simulations in Session 1, Match 4

for learning to spread throughout the network. We expect player 4 to learn more quickly, due to the stabilizing decision-making by players 5 and 6. Likewise for player 7, who is located on the other side of players 5 and 6. We expect to see learning spread from neighbor to neighbor, moving away from the two free-riders in each direction, until we reach player 13, who is farthest removed from players 5 and 6. Figure 4.17 illustrates the configuration and the hypothesized spread of learning around the network.

In Figure 4.18, we show that this pattern of learning contagion is reflected in both the simulations and the data. Players 4 and 7 converge to zero contributions within two or three rounds. Player 8

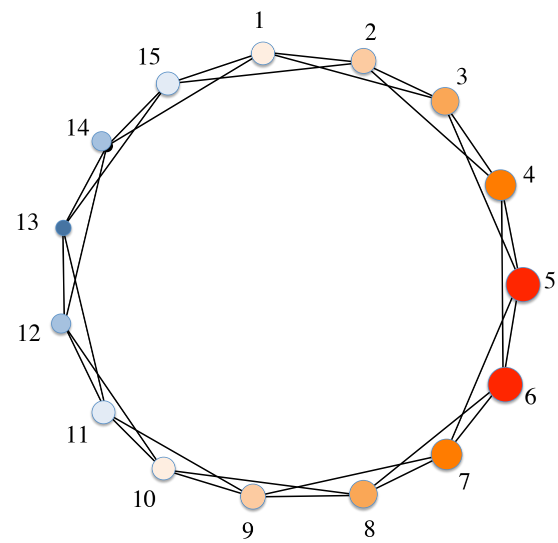


Figure 4.17: Configuration in Session 2, Match 4

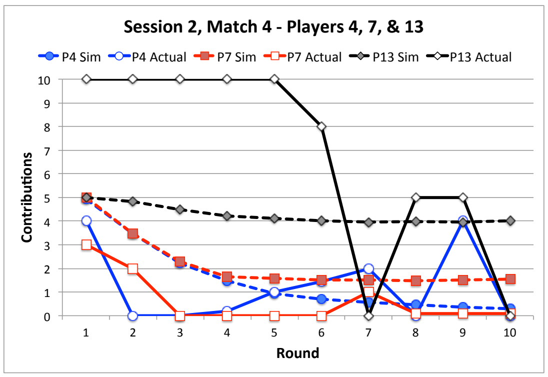
is also quick to begin playing the Nash equilibrium contribution, although player 3 seems to learn more slowly. In turn, player 8's neighbors, player 9 and player 10, learn considerably faster than players 1 and 2, who are neighbors to player 3. Finally, for player 11, 12, 14, and 15, contributions only start to approach zero in the later rounds of the match. This exhibition of contagious learning is particularly stark in the data, and especially well-portrayed by the differences in learning as we move around the circle in opposite directions.

Session 6, Match 4. In this match, there are four altruists (at nodes 2, 4, 6, and 7), and one free-rider (at node 10). The Nash equilibrium contributions for the conditional cooperators are

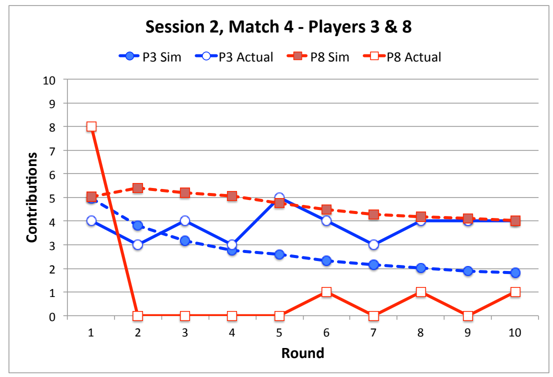
$$y_1^* = 7.76, \quad y_3^* = 10, \quad y_5^* = 10, \quad y_8^* = 6.12, \quad y_9^* = 3.21$$

$$y_{11}^* = 0, \quad y_{12}^* = 0, \quad y_{13}^* = 0, \quad y_{14}^* = 2.24, \quad y_{15}^* = 5.$$

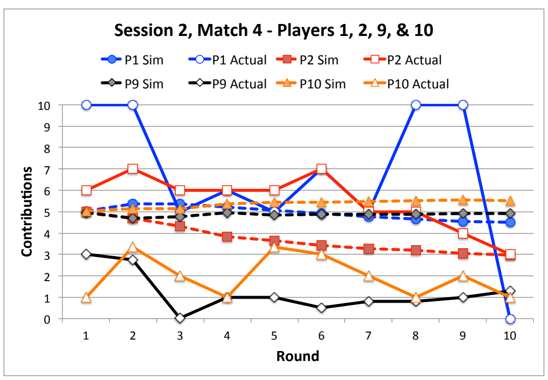
In this match, our predictions about learning contagion are somewhat mixed compared with the previous matches. On the one hand, players 3 and 5 remain several tokens below their Nash equilibrium contributions, despite being surrounded by several pure altruists. Player 8 seems to be on the right track, until player 9, who doesn't learn as quickly as we might expect, drops his contribution



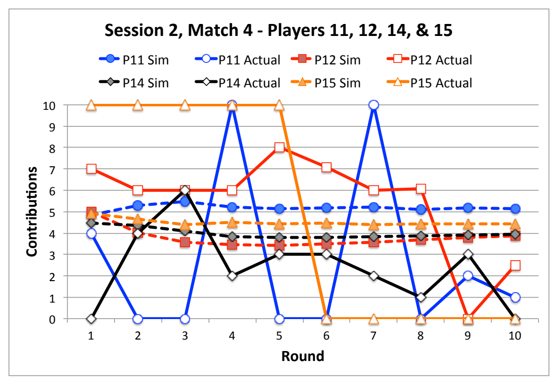
(a) Players 4, 7, & 13



(b) Players 3 & 8



(c) Players 1, 2, 9, & 10



(d) Players 11, 12, 14, & 15

Figure 4.18: Experimental data and learning simulations in Session 2, Match 4

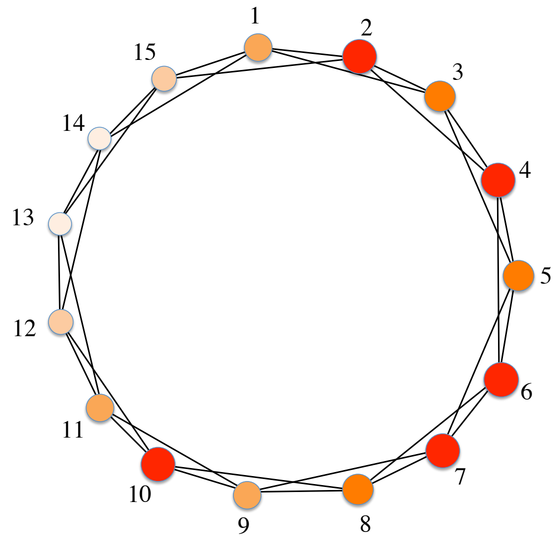


Figure 4.19: Configuration in Session 6, Match 4

from 10 to 0, after which player 8 also diverges away from his equilibrium contribution. On the other hand, some of the players who we expect to take longer at learning their equilibrium contributions actually learn quickly. Player 11 (perhaps also aided by player 9's zero contributions) converges to 0 (his equilibrium level) by round 6, and consequently, so does player 12, whose learning relies partially on how quickly player 11 learns. The contagion effect seems to be strong in this part of the network, so much so that even players 13 and 14, who we predict to be the slowest to converge to their equilibrium contributions, are close by the end of the ten rounds. Thus, although the predictions are not met with respect to some of the more likely players, there is considerable support for contagious learning on the other side of the network.

Evidence from the Core-Periphery Networks

Session 5, Match 4. Consider the configuration shown in Figure 4.21, which represents Session 5, Match 4. There is a single free-rider, player 3, who is in the periphery of the network. In this setting, there are two Nash equilibria. In one, every player contributes zero. In the other, every player in the periphery contributes 0, while every player in the core contributes the full endowment of 10 tokens. We ran learning simulations for this configuration over 100 periods and found that players' contributions converge to the latter equilibrium, with core players contributing 10 tokens. Likewise in the 10-period simulations, average contributions for the core players trend upwards, reaching approximately 7 tokens for players 6, 7, and 8, and 6 tokens for player 9 and 10. This highlights an interesting twist on the type of learning contagion observed for circles, connected cliques, and two-step circle networks. In this network, player 3 acts as an anchor which *prevents* players 9 and 10 from learning, rather than enhance their learning.

The learning predictions in this case are not as clear as in other networks. First, we predict that the learning will spread around the periphery from player 3 to his immediate peripheral neighbors, players 2 and 4, then on to players 1 and 5. At the same time, all the players in the core, since they are conditional cooperators, will attempt to equalize their own payoff with the average payoff of their neighborhood. Since they have more neighbors than the peripheral players, their payoffs are higher than the peripheral players' payoffs, even if those players are free-riding. As a result, we expect the learning dynamics to drive the contributions up for the core players, until they reach the full endowment. In this configuration, the core players who are not connected to the free-rider (player 3) should begin to learn this optimal strategy relatively fast, while those who are connected to player 3 may be slowed by the lower contributions. As players 2 and 5, followed by player 1 learn to play their zero contributions Nash equilibrium strategy, there may be some initial resistance from the players in the core, although all of the players would eventually converge to their Nash equilibrium contributions if they played long enough.

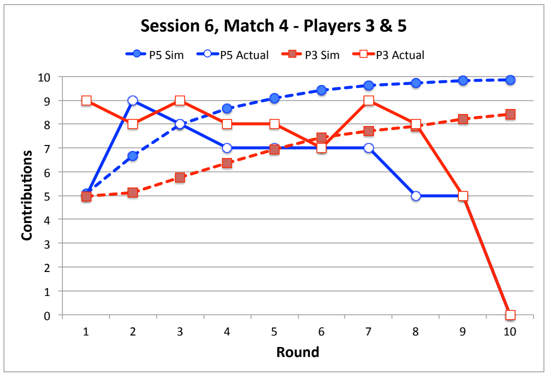
We plot the actual contributions against the simulated contributions in Figure 4.22.

4.6 Conclusion

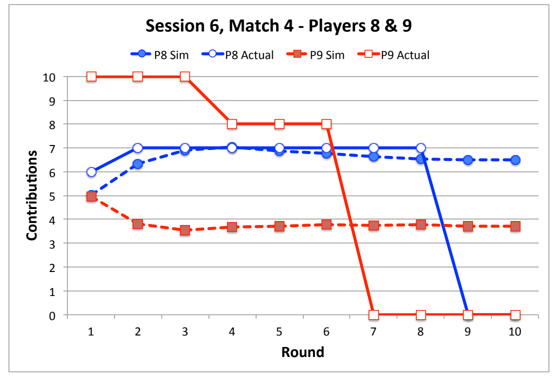
Our main finding in this chapter is that behavior in a range of different network structures is consistent with the hypothesis of contagious learning. We provide extensive evidence at the individual match level that learning spreads from the anchoring contributions of free-riders and altruists, through the network via their neighbors and their neighbors' neighbors. This evidence reinforces the result obtained in Chapter 3 for the circle network. We hope that this finding will motivate further research into the relationship between learning and network structure, both in public goods experiments and in other experimental games.

We also have examined contributions in groups consisting only of conditional cooperators, finding that average contributions typically decline over time, to a level below half of the endowment. We conjecture that this may be related to the learning dynamics, and that if so, learning dynamics may provide a means of selecting an equilibrium from among the multiple that exist for groups with only conditional cooperators. This conjecture is another avenue for future research into the behavior of individuals in repeated network public goods games.

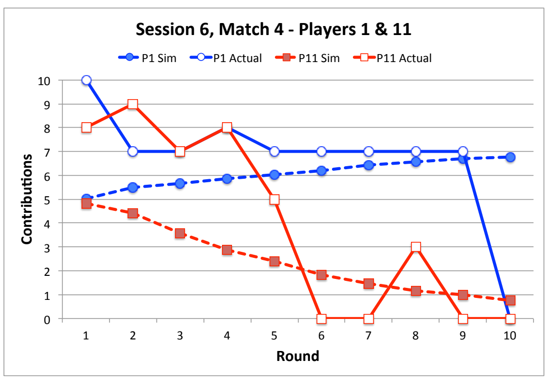
Finally, we do also observe subject behavior which is inconsistent with the predictions of the other-regarding preferences component of our behavioral model. This makes it difficult to compare average contributions, or average public good consumption across networks, since differences in the contributions may reflect a different composition of free-riders, altruists, and conditional cooperators, or even a different set of configurations. Still, with more data, and perhaps with a little more control over the precise conditions, compositions, and configurations across networks, we may be able to make more meaningful comparisons and discover more about the relationship between network structure and contributions.



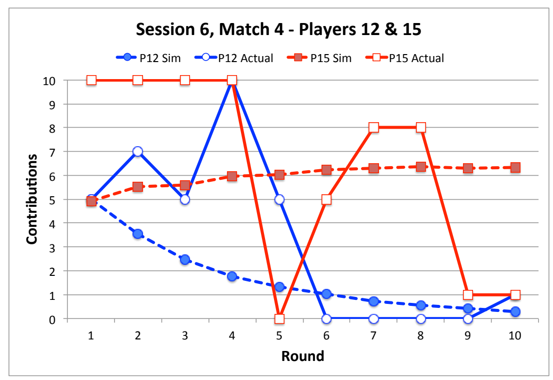
(a) Players 3 & 5



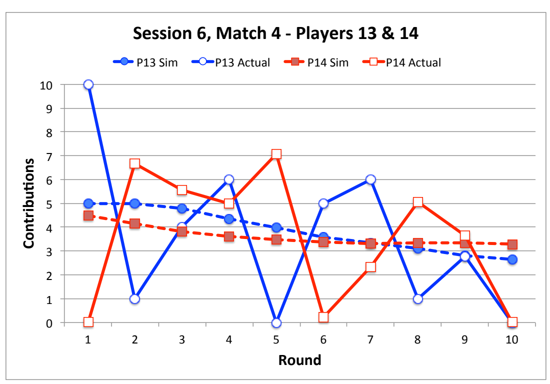
(b) Players 8 & 9



(c) Players 1 & 11



(d) Players 12 & 15



(e) Players 13 & 14

Figure 4.20: Experimental data and learning simulations in Session 6, Match 4

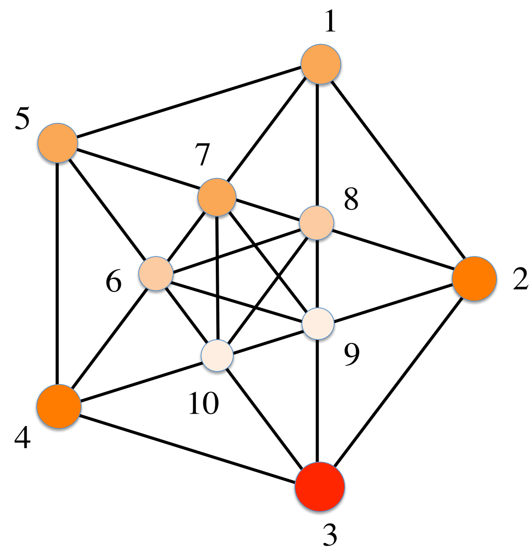
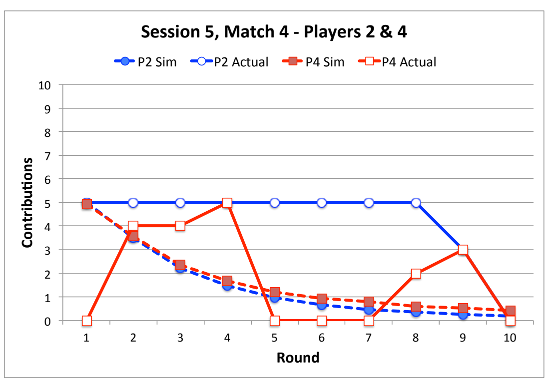
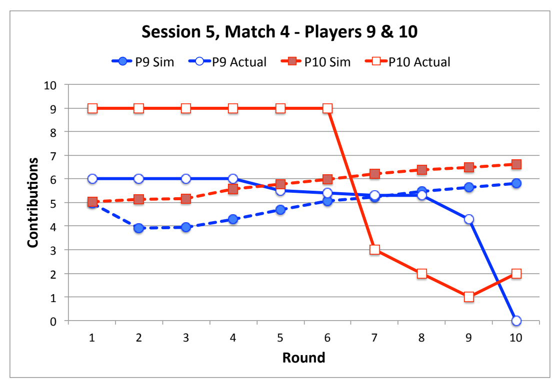


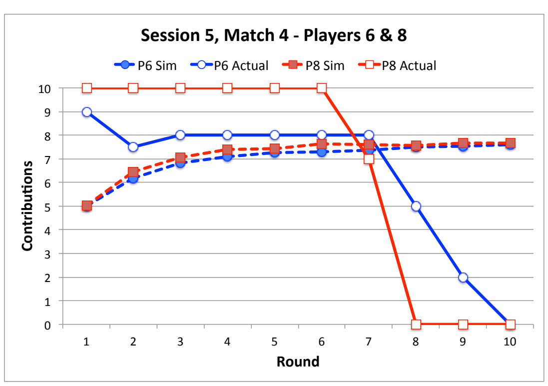
Figure 4.21: Configuration in Session 5, Match 4



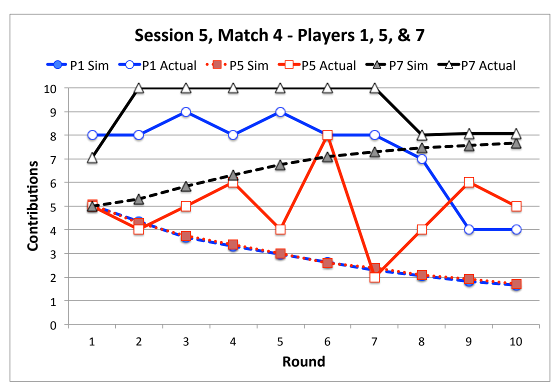
(a) Players 2 & 4



(b) Players 9 & 10



(c) Players 6 & 8



(d) Players 1, 5, & 7

Figure 4.22: Experimental data and learning simulations in Session 5, Match 4

Appendices

Appendix A

Proofs of Chapter 2

A.1 Proof of Proposition 3.4.1

The equilibrium strategies are derived from the players' marginal utility functions. Let $\mathbf{y}$ be the profile of contributions and suppose that y_i is less than the average contribution made by player i's reference group,

$$y_i < \sum_{j \in R_i} \frac{y_j}{|R_i|}. \tag{A.1}$$

Then player i's marginal utility from increasing y_i is

$$\frac{\beta^i}{|R_i|} - (1 - A). \tag{A.2}$$

On the other hand, if y_i is greater than the average contribution made by player i's reference group,

$$y_i > \sum_{j \in R_i} \frac{y_j}{|R_i|}, \tag{A.3}$$

then player i's marginal utility from increasing y_i is

$$\frac{\beta^i + \gamma^i - (1 - A + \gamma^i)|R_i|}{|R_i|}. \tag{A.4}$$

Since $\gamma^i > 0$ and $|R_i| \geq 2$, the expression in (A.2) is strictly greater than the expression in (A.4). Thus, for any profile of contributions $\mathbf{y}$, player i's marginal utility of her own contribution is nonnegative whenever the expression in (A.4) is nonnegative, i.e., when

$$\beta^i \geq (1 - A + \gamma^i)|R_i| - \gamma^i, \tag{A.5}$$

and nonpositive whenever the expression in (A.2) is nonpositive, i.e., when

$$\beta^i \leq (1 - A)|R_i|. \tag{A.6}$$

A player who does not satisfy either of the inequalities in (A.5) or (A.6) is a conditional cooperator, for whom marginal utility is positive when her contribution is lower than the average contribution in her reference group, and negative when her contribution is more than the average in her reference group. Thus, for each profile of others' contributions, her best response is to contribute the average amount contribute by the other players in her reference group.

The Nash equilibrium of a given realization of this game is solved in the following manner. First, classify the altruists and free-riders, whose equilibrium behavior is unconditional. Then find the equilibrium contributions of the conditional cooperators as the solution to a system of linearly independent equalities. $\square$

A.2 Proof of Proposition 3.4.2

Let $\mathbf{y}$ be the profile of contributions and suppose that π_i is greater than the average payoff in player i's reference group,

$$\pi_i \geq \sum_{j \in R_i} \frac{\pi_j}{|R_i|}. \tag{A.7}$$

Then player i's marginal utility from increasing y_i is

$$\frac{|N_i(G) \cap R_i|}{|R_i|} A\beta^i - (1-A)\left(1 + \frac{\beta^i}{|R_i|}\right). \tag{A.8}$$

On the other hand, if π_i is less than the average payoff in player i's reference group,

$$\pi_i \leq \sum_{j \in R_i} \frac{\pi_j}{|R_i|}, \tag{A.9}$$

then player i's marginal utility from increasing y_i is

$$\frac{|N_i(G) \cap R_i|}{|R_i|} A(\beta^i - \gamma^i) - (1-A)(1 + \gamma^i + \frac{\beta^i - \gamma^i}{|R_i|}). \tag{A.10}$$

As with the previous case, since $\gamma^i > 0$ and $|R_i| \geq 2$, the expression in (A.8) is strictly greater than the expression in (A.10). Thus, for any profile of contributions $\mathbf{y}$, player i's marginal utility of her own contribution is nonnegative whenever the expression in (A.10) is nonnegative, i.e., when

$$\beta^i \geq \gamma^i + \frac{(1+\gamma^i)(1-A)|R_i|}{A\left(|N_i(G) \cap R_i| + 1\right) - 1}, \tag{A.11}$$

and nonpositive whenever the expression in (A.8) is nonpositive, i.e., when

$$\beta^i \leq \frac{(1-A)|R_i|}{A\left(|N_i(G) \cup R_i| + 1\right) - 1}. \tag{A.12}$$

These two inequalities can be interpreted in the same way as for the variation of the game in which players use average contribution as their reference statistic. First, if player i is sufficiently altruistic, then her best response to any profile of others' contributions is to contribute everything. On the other hand, if player i is self-interested enough, then her best response is always to contribute nothing. The first player is an altruist, and the second is a free-rider.

A player who does not satisfy either of the inequalities in (A.11) or (A.12) is a conditional cooperator, for whom marginal utility is positive when her payoff is greater than the average payoff in her reference group, and negative when her payoff is less than the average in her reference group. Thus, for each profile of others' contributions, her best response is to contribute an amount that equalizes her payoff with the average payoff earned by the players in her reference group. That is, if

$$\frac{(1-A)|R_i|}{A\Big(|N_i(G)\cap R_i|+1\Big)-1} < \beta^i < \frac{(1+\gamma^i)(1-A)|R_i|}{(1-\gamma^i)\Big[A\Big(|N_i(G)\cap R_i|+1\Big)-1\Big]}, \tag{A.13}$$

then player i's best response to $\mathbf{y}_{-i}$ is to choose y_i^* such that

$$\pi_i(G,(y_i^*,\mathbf{y}_{-i})) \quad = \quad \sum_{j\in R_i}\frac{\pi_j(G,(y_i^*,\mathbf{y}_{-i}))}{|R_i|}. \tag{A.14}$$

The Nash equilibrium of a given realization of this game can be solved for using similar steps as for the first variation of the game, with one important difference. First, classify the altruists and free-riders, whose equilibrium behavior is unconditional. Then find the equilibrium contributions of the conditional cooperators as the solution to a system of linear inequalities of the form in equation (A.14), subject to the feasibility constraints on the players' actions. These additional constraints on the system of equations are the key difference between the first and second variations of the game. For some players, it may be the case that, in equilibrium, they would like to contribute a negative amount, or contribute more than their endowment, in order to equalize their payoff with the average payoff of their reference group. Thus, solving for the equilibrium behavior is slightly more difficult in this variation of the game than in the game where players care about average contribution. □

Appendix B

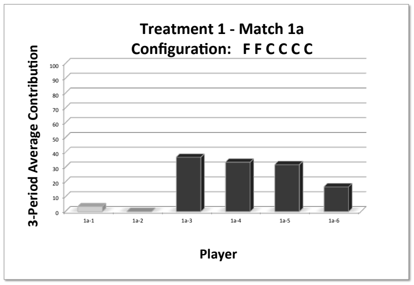
Supplementary Analysis for Chapter 3

This Appendix provides more detailed analysis of the evidence for Result 3.5.1 and Result 3.5.2. It complements the snapshot provided in Section 3.5. The analysis is broken down by information treatment for ease of comparison.

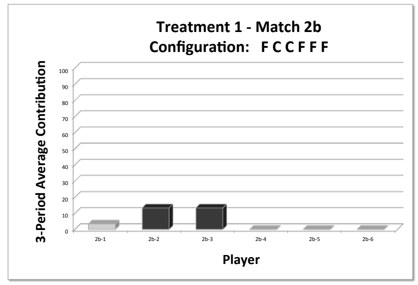
Treatment 1

In Figure B.1, I reproduce the 3-period average contributions of each player for a subset of the individual matches played under Treatment 1. Players who are conditional cooperator types are shown in black, while the free-rider and altruist types are shown in light gray. In Match 1a, there are two adjacent free-riders; in 2b and 3a, there are 4 adjacent free-riders; and in Match 4b, there is just one free-rider. In all of these matches, the benchmark Nash equilibrium is for every player to contribute 0. The subfigures highlight the fact that for Matches 2b and 3a, the contributions of the conditional cooperators are lower on average than the contributions made by the conditional cooperators in Matches 1a and 4b.

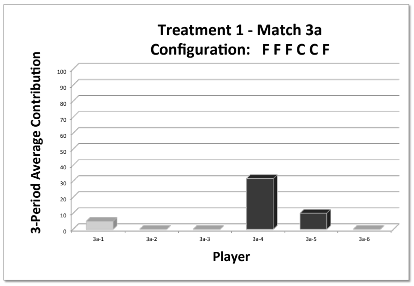
Match 4b provides support for Result 3.5.2, since the players' contributions are clearly increasing as we move further away Player 1 (the lone free-rider), peak at Player 4, who is the farthest away from the free-rider, then decrease as we move closer to the free-rider around the circle network. Still looking at Treatment 1, both Result 3.5.1 and Result 3.5.2 are also supported by Match 1b (which has 1 altruist) and Match 4a (which has 4 altruists), which are compared in Figure B.2. In



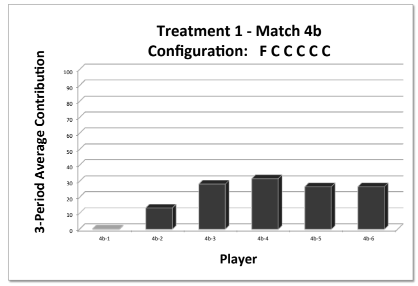
(a) Match 1a



(b) Match 2b



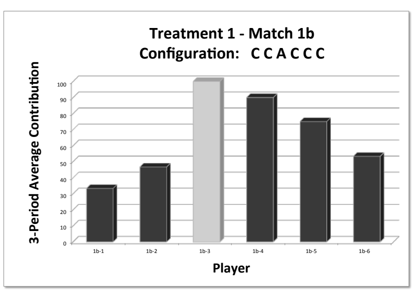
(c) Match 3a



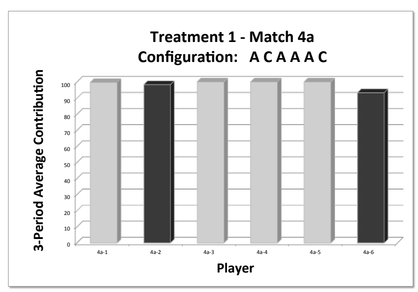
(d) Match 4b

Figure B.1: Treatment 1 matches with free-riders and conditional cooperators only

both matches, the benchmark Nash equilibrium is for all conditional cooperators to contribute their endowment of 100. As we can see, the contributions of the conditional cooperators are significantly higher in Match 4a than in Match 1b. Moreover, in Match 1b, simulation and actual contributions are declining as we move further away from Player 3, who is the lone altruist, which supports Result 3.5.2.



(a) Match 1b

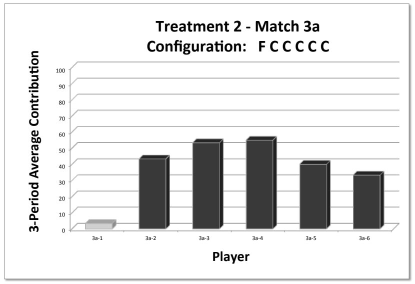


(b) Match 4a

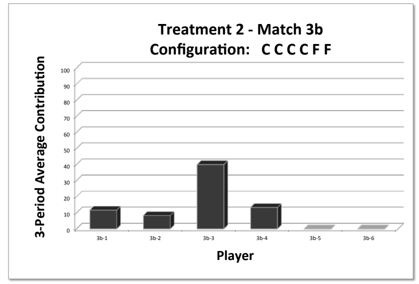
Figure B.2: Treatment 1 matches with altruists and conditional cooperators only

Treatment 2

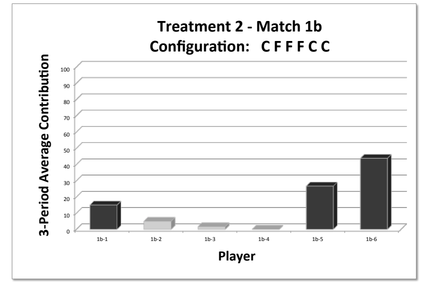
In Treatment 2, there are two matches with 1 free-rider (3a and 4b), two matches with 2 adjacent free-riders (3b and 5b), 2 matches with 3 adjacent free-riders (1b and 2a), and 1 match with 5 free-riders (2b). these matches support Result 3.5.1, and all except for Match 4b illustrate the dependence on relative positions summarized in Result 3.5.2. Comparing the left column with the middle column and the right column in Figure B.3, gives a good indication of how, even though all of these matches have the same benchmark Nash equilibrium (zero contributions by everyone), the contributions of the conditional cooperators are positive and decreasing in the number of free-riders in the match. Moreover, aside from Match 4b, contributions are higher for players who are positioned further from the free-riders than for those who are closer, which supports Result 3.5.2.



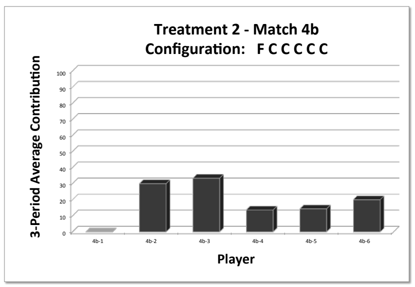
(a) Match 3a



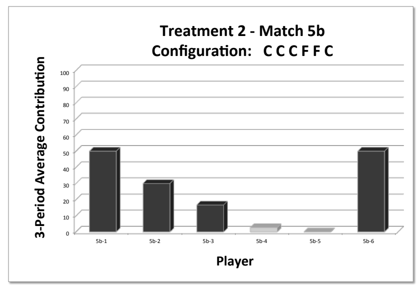
(b) Match 3b



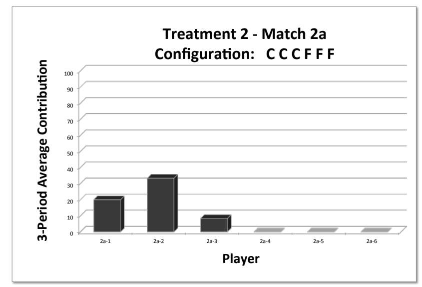
(c) Match 1b



(d) Match 4b

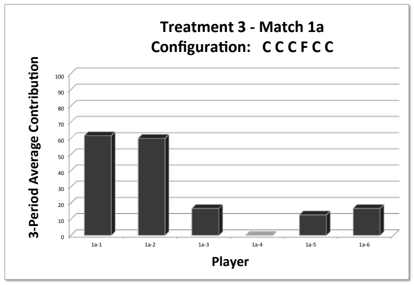


(e) Match 5b

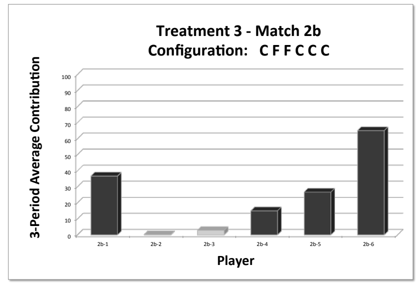


(f) Match 2a

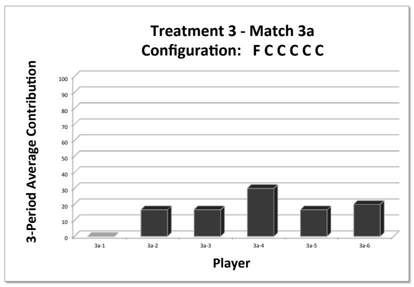
Figure B.3: Treatment 2 matches with free-riders and CCs only. The number of free-riders in the matches is increasing from left to right.



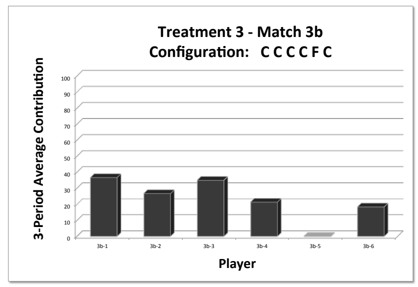
(a) Match 1a



(b) Match 2b



(c) Match 3a



(d) Match 3b

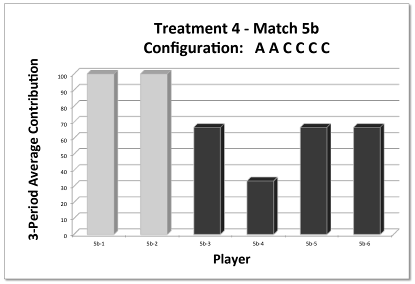
Figure B.4: Treatment 3 matches with free-riders and CCs only

Treatment 3

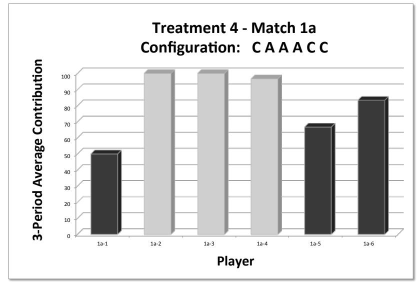
Treatment 3 provides the most remarkable evidence to support Result 3.5.2. Consider the matches displayed in Figure B.4. In Match 1a, contributions are monotonically increasing for both simulations and data, as we move further away from Player 4 who is the sole free-rider. The same is true in Match 2b as we move away from Players 2 and 3 (the free-riders), Match 3a as we move further away from Player 1, and Match 3b as we move away from Player 5.

Treatment 4

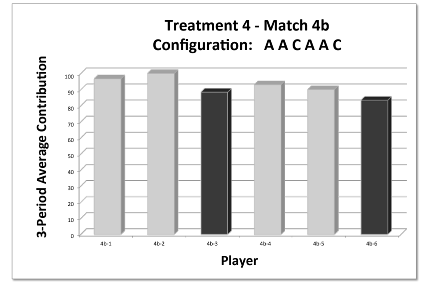
Figure B.5 reproduces Matches 5b (2 altruists), 1a (3 altruists), and 4b (4 altruists) under Treatment 4. In support of Result 3.5.1, the contributions of the conditional cooperator types are highest in Match 4b, and lowest in Match 5b. On the other hand, Figure B.6 displays four matches under Treatment 4 that exhibit support for Result 3.5.2.



(a) Match 5b

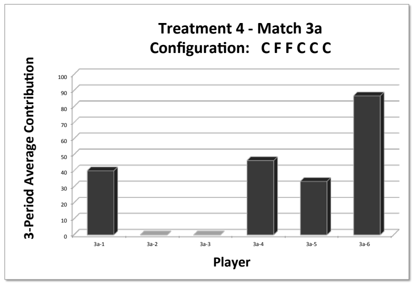


(b) Match 1a

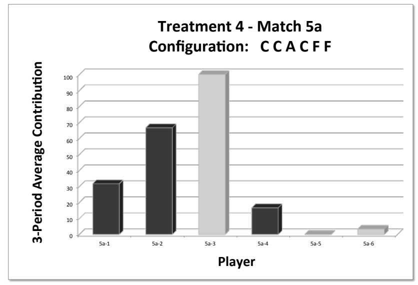


(c) Match 4b

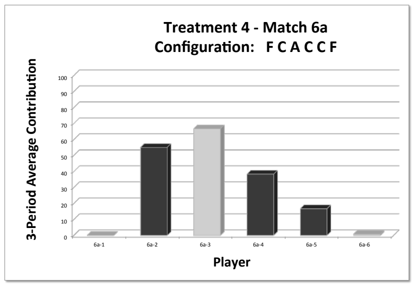
Figure B.5: Treatment 4 matches — Contributions and the # of altruists



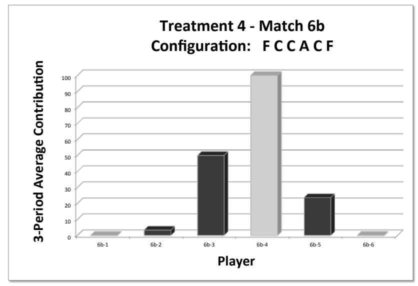
(a) Match 3a



(b) Match 5a



(c) Match 6a



(d) Match 6b

Figure B.6: Treatment 4 matches — Effects of players' relative positions

Appendix C

Experiment Instructions for Chapter 3

Instructions

Thank you for agreeing to participate in this experiment. During the experiment, please give us your full attention and follow the instructions carefully. Please turn off your cell phones, and refrain from chatting with other subjects, opening other applications on your computer, or engaging in other activities. At the end of the experiment, you will be paid discreetly in cash, based on the payoffs you earn. What you earn depends partly on your own decisions, partly on the decisions of others, and partly on chance. Do not talk or try to communicate with other participants during the experiment.

Following these instructions, there will be a practice session and a short comprehension quiz. You must answer all questions on the quiz correctly before you can continue to the paid sessions. In the experiment, your earnings will be denominated in **tokens**. At the end, these earnings will be converted to US dollars at the rate of **500 tokens** to **1 US Dollar**.

This experiment consists of four matches. In each match, there will be 15 rounds. For each match, you will be divided into TWO groups of SIX members each. You will be randomly assigned to exactly one of these two groups and you will not know who out of the other participants is in your group. Do not let any other subject know which group you are in. You will remain in this group for the entire first match. For each other match, you will be randomly rematched into two (possibly different) groups of SIX members each. Thus, your group will be fixed during a given match, but may be different across matches. Other than the set of players in each group, the parameters and the features of the match will be the same for both groups.

Each match will proceed as follows. At the start of the match, you will be randomly assigned to a position (node) in the network depicted in Figure C.1. The other 5 members of your group will be assigned to the other positions, so that only one member is at any position, and all positions are filled.

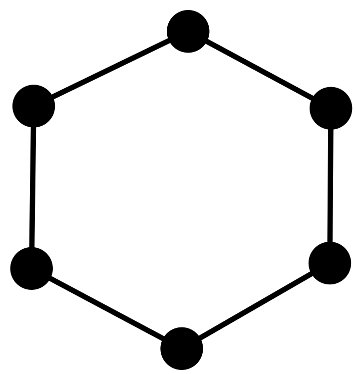


Figure C.1: Circle network with 6 agents

The match will consist of 15 rounds. During the match, your position will be identified by a node labeled “You” and a player number. Your player number and your location will remain fixed throughout each round of the match. Likewise, your group members, and their locations will be fixed throughout each round of the match. If your node is connected to another node, then that node will be displayed in red to indicate the connection. The players located at the red nodes are your **direct neighbors** in the network.

In each round, you will face exactly the same decision problem. At the start of every round, you will be given an endowment of 100 **tokens**. You must decide how much of this endowment to contribute to a given project, and how much to keep for yourself. You cannot contribute a negative amount nor can you contribute more than **100 tokens** towards the project. You may choose any number up to two decimal places within that range. In a given round, your earnings from the project

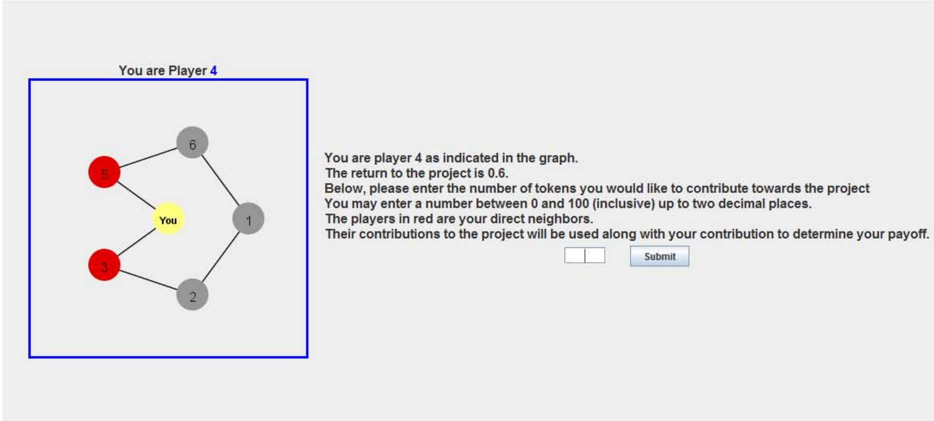


Figure C.2: The decision screen for a match

depend on your allocation to the project in that round and the allocations made by your **direct neighbors** in that round. Specifically, your earnings from the project are calculated by adding your contribution and the contributions of your **direct neighbors**, then multiplying the total by the return factor, which is 0.6. Your earnings from the project will then be added to whatever number of tokens you keep (which will be 100 minus your contribution) to give your overall payoff from the round.

For example, suppose you allocate 40 **tokens** to the project and keep 60 **tokens** for yourself, and the sum of the allocations made by your **direct neighbors** to the project is 80. Then your earnings from the project will be

$$0.6 \cdot (40 + 80) = 72,$$

while your earnings from the tokens you keep will be 60. Thus, your total earnings would be $Earnings = 72 + 60 = 132$.

To summarize, your payoffs in a given round will be equal to

$$\text{payoff} = 100 - \text{your contribution} + 0.6 \times (\text{total contributions by you and your neighbors}).$$

Each round is a separate decision problem, so your earnings in any round will depend only on

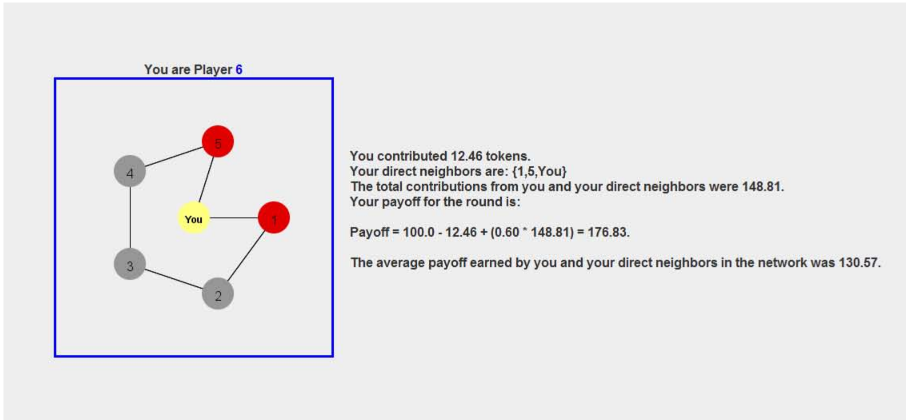


Figure C.3: The round summary screen for a match

the decisions made by you and your **direct neighbors** in that round. After each round, you will see certain information about what happened. This information will be different for each match, as will be described below and before the match. You will be given 30 seconds after each round to observe the information, then we will move to the next round. At the end of the match, all of your round payoffs will be added together to give your match payoffs. At the end of the last match, your match payoffs will be summed and converted into US dollars according to the exchange rate above.

After each round, you will see the following information in all matches.

- The amount that you contributed to the project
- The total contributions to the project from **you AND your direct neighbors**
- Your payoff from the round

In addition, match specific information will be provided as specified before each match.

Match 1

The average payoff received by **you and your direct neighbors**

Match 2

The average payoff received by **all SIX players in your group**

Match 3

The average contribution made by **you and your direct neighbors**

Match 4

The average contribution made by **all SIX players in your group**

Now we will run through a practice match with 4 rounds, so that you can familiarize yourself with the software. If you have any questions, please raise your hand. You will not be paid for this practice session. After the practice match, there will be a short quiz for you to answer, before you can proceed to the paid matches. As a reminder, please do not communicate with the other subjects in any way. Good luck.

Appendix D

Experiment Instructions for Chapter 4

Instructions

Thank you for agreeing to participate in this experiment. During the experiment, please give us your full attention and follow the instructions carefully. Please turn off your cell phones, and refrain from chatting with other subjects, opening other applications on your computer, or engaging in other activities. At the end of the experiment, you will be paid discreetly in cash, based on the payoffs you earn. How much you earn depends partly on your own decisions, partly on the decisions of others, and partly on chance. Please do not try to communicate with other participants during the experiment.

Following these instructions, there will be an unpaid practice session and a short comprehension quiz. You must answer all questions on the quiz correctly before you can continue to the paid sessions. In the experiment, your earnings will be denominated in **tokens**. At the end, these earnings will be converted to US dollars at the rate of **50 tokens** to **1 US Dollar**. The experiment consists of seven parts, called matches. Each match will be described separately before it is played.

Match 1 (unpaid practice) and Match 2

In this match, you will be divided into THREE (3) groups of FIVE (5) members each. You will be randomly assigned to one of these groups and you will not know who out of the other participants is in your group. Your group will remain the same throughout the match. The practice (unpaid) match will consist of 3 rounds, followed by a short comprehension quiz. The paid Match 2 will consist of 10 rounds. In each round, you will face exactly the same decision problem.

The Decision Problem

At the start of every round, you will be given an endowment of 10 **tokens**. You must decide how much of this endowment to contribute to a given project, and how much to keep for yourself. You may contribute any number, up to two decimal places, between 0 and 10, inclusive. In a given round, your earnings will be the sum of your earnings from the project and the number of tokens you keep for yourself. Your earnings from the project are calculated by taking the total number of tokens contributed to the project by the members of your group, and multiplying it by a return factor of 0.4. These earnings from the project will be added to the tokens you keep for yourself to give your overall payoff from the round.

For example, suppose you allocate 6 **tokens** to the project, keeping 4 **tokens** for yourself, and suppose the sum of the allocations made by the other members of your group is 14 tokens. Your earnings from the project will be

$$0.4 \cdot (6 + 14) = 8.$$

Your overall earnings will be the 4 tokens you keep for yourself plus the 8 tokens you earn from the project, for a total of 12 tokens. Each round is a separate decision problem, so your earnings in any round depend only on the decisions made by you and your group members in that round.

After each round, you will be shown

- how many tokens you allocated to the project,

- the total number of tokens allocated to the project in your group (including yours),
- your payoff from the round,
- the average payoff earned by the members of your group (including you).

Matches 3–8

In these matches, you will be randomly assigned to a position in a particular network graph. These are shown below. The decision problem will be the same for each match, although the network and your neighbors (the players to whom you are connected) will change across matches. Your neighbors in the network will be the participants assigned to positions (or nodes) that are linked by an edge to your position (node). All positions are fixed throughout a given match. Each match will consist of 10 rounds. In each round, you will face exactly the same decision problem.

The Decision Problem

At the start of every round, you will be given an endowment of 10 **tokens**. You must decide how much of this endowment to contribute to a given project, and how much to keep for yourself. You may contribute any number, up to two decimal places, between 0 and 10, inclusive. In a given round, your earnings from the project are calculated by taking the total number of tokens contributed to the project by you and your neighbors in the network, and multiplying it by a return factor of 0.4. Your earnings from the project will be added to the tokens you keep for yourself to give your overall payoff from the round.

For example, suppose you allocate 7 **tokens** to the project, keeping 3 **tokens** for yourself, and suppose the sum of the allocations made by your neighbors in the network is 28 tokens. Your earnings from the project will be

$$0.4 \cdot (7 + 28) = 14.$$

Your overall earnings will be the 3 tokens you keep for yourself plus the 14 tokens you earn from the project, for a total of 17 tokens. Each round is a separate decision problem, so your earnings in any round depend only on the decisions made by you and your neighbors in that round.

After each round, you will be shown

- how many tokens you allocated to the project,
- the total number of tokens allocated to the project by you and your neighbors,
- your payoff from the round,
- the average payoff earned by you and your neighbors.

The networks are summarized below.

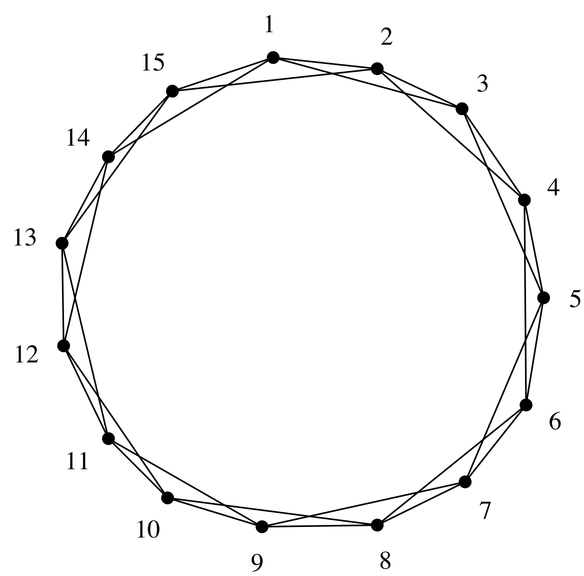


Figure D.1: Match 3 network graph

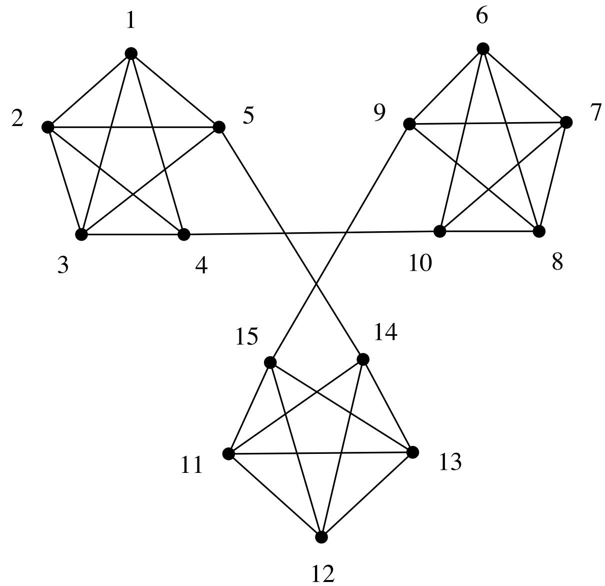


Figure D.2: Match 4 network graph

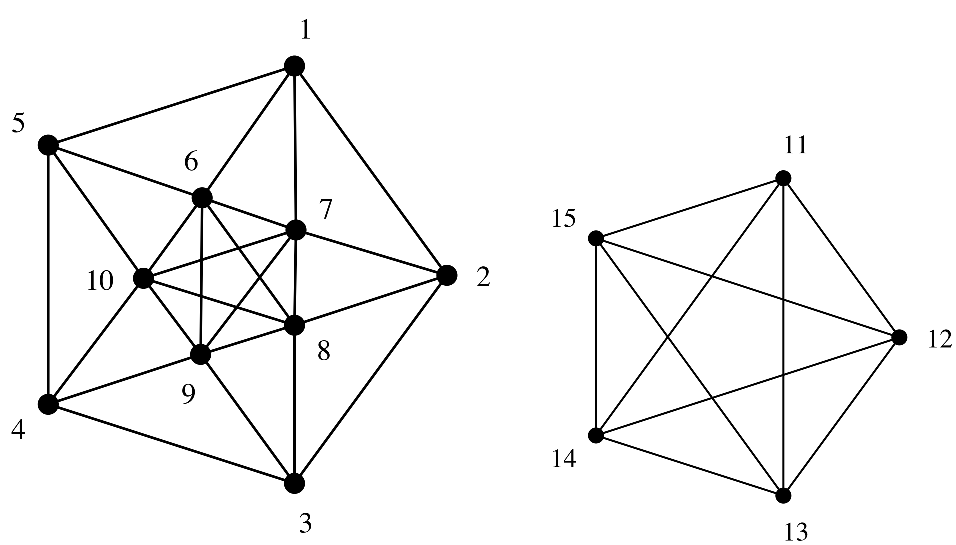


Figure D.3: Match 5 network graph

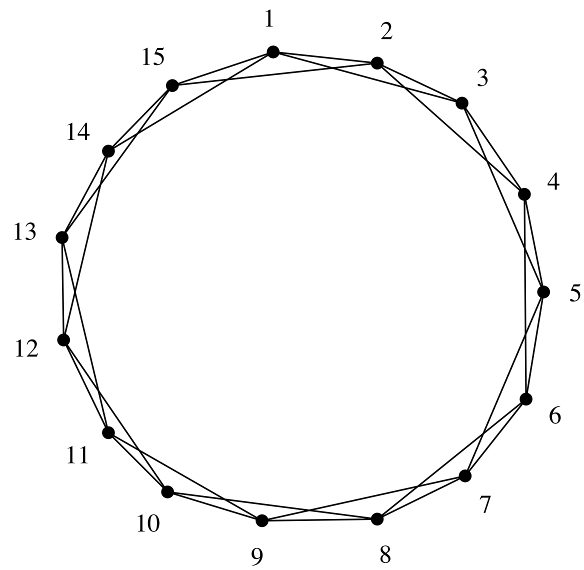


Figure D.4: Match 6 network graph

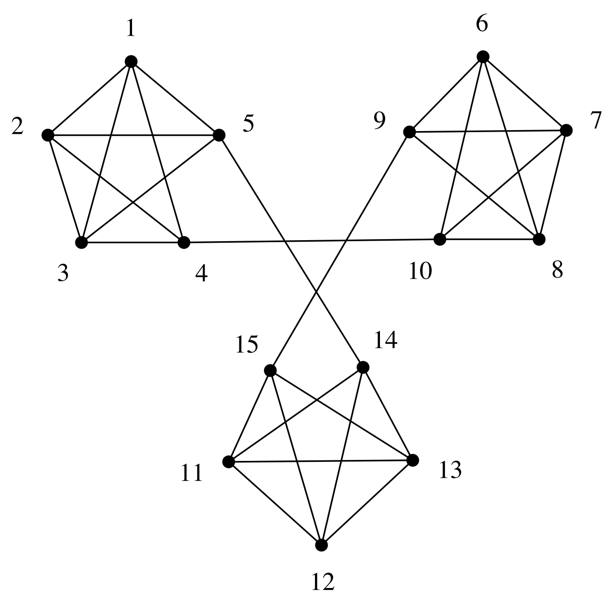


Figure D.5: Match 7 network graph

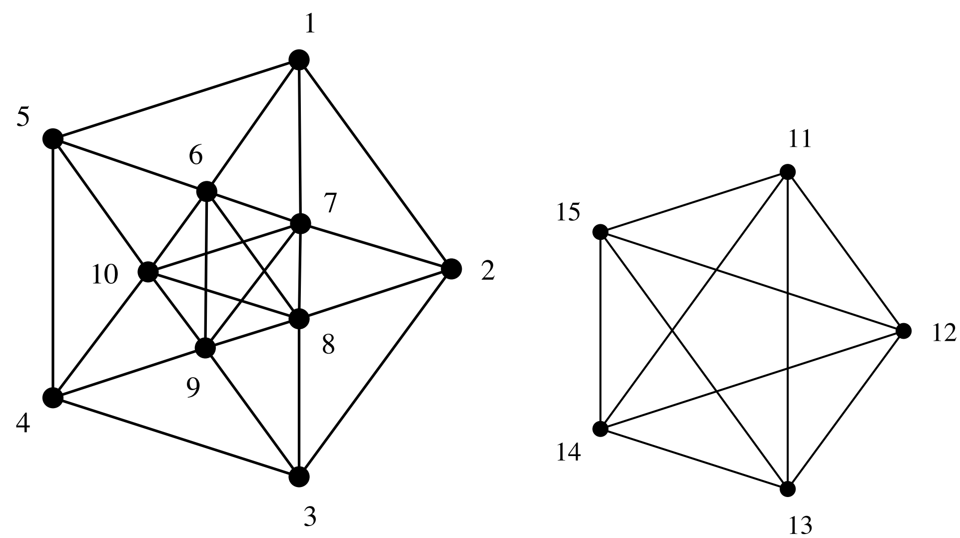


Figure D.6: Match 8 network graph

Bibliography

AMBRUS, A., AND P. PATHAK (2011): "Cooperation over finite horizons: A theory and experiments," *Journal of Public Economics*, 95(7–8), 500–512.

ARIFOVIC, J., AND J. LEDYARD (2009): "Individual Evolutionary Learning, Other-regarding Preferences, and the Voluntary Contributions Mechanism," Working paper.

BOLTON, G., AND A. OCKENFELS (2000): "ERC — A theory of equity, reciprocity and competition," *The American Economic Review*, 90, 166–193.

BRAMOULLÉ, Y., AND R. KRANTON (2007): "Public goods in networks," *Journal of Economic Theory*, 135(1), 478–494.

BULOW, J., AND J. LEVIN (2006): "Matching and Price Competition," *The American Economic Review*, 96(3), 652–668.

CAMERER, C., AND T. H. HO (1999): "Experience-weighted attraction learning in normal form games," *Econometrica*, 67(4), 827–874.

CHARNESS, G., AND M. RABIN (2002): "Understanding social preferences with simple tests," *Quarterly Journal of Economics*, 117(3), 817–869.

CHAUDHURI, A. (2011): "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 14(1), 47–83.

CRAWFORD, V. (2008): "The Flexible-Salary Match: A proposal to increase the salary flexibility of the National Resident Matching Program," *Journal of Economic Behavior & Organization*, 66(2), 149–160.

Crawford, V., and E. Knoer (1981): "Job Matching with Heterogeneous Firms and Workers," *Econometrica*, 49(2), 437–450.

Croson, R. (2007): "Theories of commitment, altruism and reciprocity: evidence from linear public goods games," *Economic Inquiry*, 45(2), 199–216.

Croson, R., E. Fatas, and T. Neugebauer (2005): "Reciprocity, matching and conditional cooperation in two public goods games," *Economics Letters*, 87(1), 95–101.

Dufwenberg, M., and G. Kirchsteiger (2004): "A theory of sequential reciprocity," *Games and Economic Behavior*, 47(2), 268–298.

Falk, A., and U. Fischbacher (2006): "A theory of reciprocity," *Games and Economic Behavior*, 54(2), 293–315.

Fehr, E., and K. Schmidt (1999): "A theory of fairness, competition and cooperation," *Quarterly Journal of Economics*, 114, 817–868.

Fischbacher, U., and S. Gachter (2010): "Social preferences, beliefs, and the dynamics of free riding in public goods experiments," *The American Economic Review*, 100(1), 541–556.

Fischbacher, U., S. Gachter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71(3), 397–404.

Frey, B., and S. Meier (2004): "Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment," *The American Economic Review*, 94(5), 1717–1722.

Gale, D., and L. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

Galeotti, A., S. Goyal, M. Jackson, F. Vega-Redondo, and L. Yariv (2010): "Network Games," *Review of Economic Studies*, 77(1), 218–244.

Gonzalez-Diaz, J., and R. Siegel (Forthcoming): "Matching and Price Competition: Beyond Symmetric Linear Costs," *International Journal of Game Theory*.

HATFIELD, J., AND P. MILGROM (2005): "Matching with Contracts," *The American Economic Review*, 95(4), 913–935.

HOPPE, H., B. MOLDOVANU, AND A. SELA (2009): "The Theory of Assortative Matching Based on Costly Signals," *Review of Economic Studies*, 76, 253–281.

ISAAC, R. M., K. MCCUE, AND C. PLOTT (1985): "Public goods provision in an experimental environment," *Journal of Public Economics*, 26, 51–74.

ISAAC, R. M., J. WALKER, AND S. THOMAS (1984): "Divergent evidence on free riding: An experimental examination of possible explanations," *Public Choice*, 43, 113–149.

KELSO, A., AND V. CRAWFORD (1982): "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 50(6), 1483–1504.

KESER, C., AND F. VAN WINDEN (2000): "Conditional cooperation and voluntary contributions to public goods," *Scandinavian Journal of Economics*, 102(1), 23–39.

KOJIMA, F. (2007): "Matching and Price Competition: Comment," *The American Economic Review*, 97(3), 1027–1031.

LEVINE, D. (1998): "Modeling altruism and spitefulness in experiments," *Review of Economic Dynamics*, 1, 593–622.

NIEDERLE, M. (2007): "Competitive Wages in a Match with Ordered Contracts," *The American Economic Review*, 97(5), 1957–1969.

RABIN, M. (1993): "Incorporating fairness into game theory and economics," *The American Economic Review*, 83(5), 1281–1302.

ROTH, A. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7(4), 617–628.

ROTH, A., AND I. EREV (1995): "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term," *Games and Economic Behavior*, 8(1), 164–212.

Roth, A., and M. Sotomayor (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, Cambridge.

Shapley, L., and M. Shubik (1972): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111–130.