Essays on Cooperation and Coordination

Thesis by

Julian Romero

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2010

(Defended May 12, 2010)

Acknowledgements

I am grateful for the help and encouragement of many people during my time at Caltech. My main advisor, John Ledyard, was always willing to talk and provided me with excellent guidance throughout my graduate career. My other committee members Jacob Goeree and Leeat Yariv were also instrumental in my time here. I would also like to thank Jean-Laurent Rosenthal for always taking the extra step to ensure success of the graduate students. My undergraduate advisor Charles Zheng for encouragement and guidance in applying and choosing a graduate school. Also, Eric Bax who served as an excellent mentor during my time working at Yahoo!.

I would also like to thank my classmates who helped make my time at Caltech more enjoyable. My classmates Eileen Chou, Umberto Garfagnini, Gabriel Katz, Bong Chan Koh, and Margaret McConnell who helped me get through the first-year. My office mates Dustin Beckett, Guido Maretto, and Mohamed Mostagir who put up with me as an office mate. Alex Brown, Laurent Mathevet, and Noah Myung who provided me with guidance on the job market. It also helped to have the support of my fellow job market candidates Sera Linardi, Yuanchen Lien, Aniol Llorente-Saguer while I was on the job market. In addition to my classmates, I would like to thank the graduate secretary Laurel Auchampaugh for all her work during my time at Caltech starting from my visit as a prospective student all the way to my time on the job market.

Finally, I would like to thank my parents, Louis and Shelley for providing me support and encouragement throughout my life. My brothers, Alex and Philip, who are also graduate students at Caltech, made my time here much better. My girlfriend, Lauren (I'm not sure what her last name is), for always distracting me from my work when I needed a break. Chrystal and Jim Johnson for making Pasadena a place that I could call home.

Abstract

This thesis examines questions related to game theory, and in particular cooperation and coordination among economic agents.

In the first chapter (joint with Noah Myung) we propose a decision making process meant to mimic human behavior. This process is implemented with computational agents. We use these computational agents to run simulations of two coordination games, the minimum-effort coordination game and the Battle of the Sexes game. We find that the computational agents exhibit behavior similar to human subjects from previous experimental work. We then use the computational testbed to develop experimental hypotheses, which are then confirmed in the laboratory using human subjects. In particular, we show that higher cost may actually lead to higher average payoffs in the minimum-effort coordination game.

The second chapter examines a model of infinitely repeated games in which agents are boundedly rational. I show that the number of equilibrium outcomes is smaller when agents are boundedly rational. Importantly, cooperative outcomes are still possible in equilibrium, even when players cannot use sophisticated strategies and are not able to perfectly monitor their opponents. The strategy that leads to cooperation is called "Win-Stay, Lose-Shift". Using this strategy, I show that cooperation is possible in equilibrium for a large class of 2×2 games. I also give necessary and sufficient conditions on equilibrium structure for $N \times 2$ games. These conditions suggest that in equilibrium, players must be able to cooperate without getting caught in long periods of conflict.

The final chapter focuses on a class of minimum-effort coordination games. I show that the symmetric quantal response equilibrium correspondence takes the shape of an s-shaped curve as long as players are sufficiently rational (sufficiently high λ). Under certain assumptions, this s-shaped

correspondence leads to hysteresis. Based on these theoretical results, I develop experiments with the minimum-effort coordination game, and test the hysteresis hypothesis in the laboratory. I find evidence that this hysteresis does occur when human subjects play the minimum-effort coordination game in the lab.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation is divided into three chapters and focuses on questions related to game theory and in particular repeated interactions between economic agents. The goal of game theory is to provide predictions about how individuals act in different games. In certain situations, there are incentives for one or more of the agents to cheat and gain an upper hand on their opponent. Other times, there is no incentive to cheat, but the agents still must somehow coordinate their actions to attain optimal outcomes. The goal of this dissertation is to better understand how these agents are able to cooperate and coordinate in these situations.

A combination of theory, computation and experiments are used in this dissertation. Each of these tools has unique benefits. Economic theory allows us to focus on a specific problem and prove general theorems that hold in all situations that the assumptions allow. Experiments allow us to examine human behavior in a controlled environment, allowing us to focus on the specific aspect of the decision making process in which we are interested. Computation allows us to run cheap simulations of models that are difficult to grasp analytically. The combination of the three areas provides a solid framework for better understanding the human decision making process in these repeated interactions.

The first chapter, "Learning to Coordinate Through Pattern Recognition" (joint with Noah Myung), focuses on the computational aspect of learning in repeated games. Economists have chosen two main techniques to study human interactions in repeated games. Economic theorists have presented learning algorithms and examined the long term convergence properties of these algorithms. Experimental economists have used human subjects in a controlled environment to help better understand these interactions. While much progress has been made in both of these areas, there is somewhat of a disconnect between the two. The theoretical learning models which have guaranteed long run convergence typically don't match the human subjects in experiments who typically converge to equilibria in games that are only repeated a relatively small number of periods.

In this chapter, we develop a computational learning algorithm that is used as an experimental testbed. The computational agents make their choices in a boundedly rational manner. They randomly sample other actions to determine what to play, they group similar payoffs together to determine what is their best action, and they make choices from a distribution. A key component to this chapter is that the agents use pattern recognition to make predictions about the actions of their opponents. If their predictions are accurate, they make more accurate choices.

This computational learning algorithm is used to create a testbed that is used to run simulations of games. The simulation data matches experimental data for both minimum-effort coordination games and the Battle of the Sexes game. In particular, the computational agents are able to learn to alternate between two equilibria in the Battle of the Sexes game, an outcome that has been observed experimentally. One benefit of using this computational testbed is that it allows us to run a large number of experimental trials over a multitude of parameter values at virtually no cost. These simulations can then be used to develop experimental hypotheses which can be tested in the lab. So, we run simulations of the minimum-effort coordination game over a wide range of cost values. The simulations suggest that higher costs in minimum-effort coordination games lead to faster rates of convergence. Therefore when players face higher costs, there are opposing effects. On one hand, higher costs decrease payoffs for any fixed strategy profile and often lead to lower effort levels. However, faster convergence means fewer round of inefficient non-convergence. The simulations suggest that the increase in payoffs due to faster convergence actually outweighs the cost due to lower effort levels. Using the simulations to design experiments, we test these hypotheses in the laboratory, and find that the hypotheses are confirmed.

The second chapter, "Bounded Rationality in Repeated Games," examines a model of infinitely repeated games in which agents are boundedly rational. When full rationality is assumed in repeated interactions, a plethora of equilibrium outcomes emerge. This large set of outcomes makes it difficult to predict how individuals will behave when faced with these situations. Ideally, a model would have a small number of outcomes which are also verified by experimental data. Many of the strategies used to construct the equilibria in the full rationality model require that players have a large memory and don't make mistakes. It is possible that individuals are boundedly rational, and therefore are not able to implement some of these complicated strategies from the full rationality case. After all, models that assume bounded rationality often lead to sharper predictions about real world outcomes than their full rationality counterparts. Therefore, this chapter asks, does the introduction of boundedly rational agents lead to a smaller set of outcomes in equilibrium?

I present a model of infinitely repeated games in which players have a bound on their memory, in particular they use strategies that can be represented by finite automata. In addition, players are not able to perfectly monitor the actions of their opponents. Using this model, I show that the number of equilibrium outcomes is smaller when agents are boundedly rational. Importantly, cooperative outcomes are still possible in equilibrium, even when players can't use sophisticated strategies and are not able to perfectly monitor their opponents. The strategy that leads to cooperation is called "Win-Stay, Lose-Shift". When players use this strategy, I show that cooperation is possible in equilibrium for a large class of 2×2 games. I also use the intuition from WSLS to get necessary and sufficient conditions on equilibrium structure for $N \times 2$ games. These conditions suggest that in equilibrium, players must be able to cooperate without getting caught in long periods of conflict.

The third chapter, "Hysteresis in Coordination Games," examines equilibrium selection in coordination games. When games have multiple equilibria, it may be difficult to predict which equilibrium the players are going to play. There has been much work focusing on refining the set of equilibria to allow for sharper predictions when the game has multiple equilibrium. In this paper, we examine the effect of hysteresis on equilibrium selection. A system exhibits hysteresis if the outcome of a game depends on the history leading up to that game. For example, If the cost starts low then increases to the middle, one equilibrium is selected, while if the cost start high and then decreases to the middle, another equilibrium is selected. I show that under certain conditions this hysteresis is possible in coordination games.

This chapter focuses on a class of minimum-effort coordination games. I show that the symmetric quantal response equilibrium correspondence takes the shape of an s-shaped curve as long as players are sufficiently rational (sufficiently high λ). Under certain assumptions, this s-shaped correspondence leads to hysteresis. Based on these results, I develop experiments with the minimum-effort coordination game, and test the hysteresis hypothesis in the laboratory. The experiments provide evidence that hysteresis does occur when human subjects play the minimum-effort coordination game in the lab.

One important implication of this hysteresis is that small changes in the cost parameter can lead to the selection of a different equilibrium. This result suggests that a temporary change in the payoff could be used as a means of moving from a bad equilibrium to a better equilibrium. For example, a city might offer a temporary subsidy for public transportation for one year, and then change the price back to the original value. It is possible that once people have switched their behavior, they will remain there even when the price is increased again.

Chapter 2

Computational Testbeds for Coordination Games

2.1 Introduction

Economists have chosen two avenues to understand human interactions in repeated games. The theorists have focused on presenting learning algorithms and proving results on the convergence properties of these algorithms. Experimentalists have used human subjects in laboratory environments to better understand behavior in these interactions. In this paper, we develop a third avenue, computation, which complements both the theoretical and experimental literature. In particular, we build computational agents which exhibit behavior consistent with the behavior of human subjects in two repeated coordination games. Using this computational testbed, we run simulations of experiments, develop testable hypotheses, and then verify these hypotheses with human subjects in the experimental laboratory.

We focus on coordination games, because by definition they must have multiple equilibria. Because of this multiplicity of equilibria, many experiments have been run on these games to examine the equilibrium selection process. Based on these experiments, some common behaviors have emerged which have been observed repeatedly in the experimental lab. In one game, the minimum-effort coordination game, the players have identical interests. In experiments with minimum-effort coordination games, Van Huyck, Battalio, and Beil (1990) and Goeree and Holt (2005) show that subjects typically converge to a Nash equilibrium, but which Nash equilibrium depends on the number of subjects and the risk involved. In the other game, the Battle of the Sexes game, the players have opposing interests. Often in the Battle of the Sexes game, subjects learn to alternate between the two equilibrium outcomes as exhibited in Rapoport, Guyer, and Gordon (1976) and McKelvey and Palfrey (2001).

The computational agents' decisions are made based on a learning algorithm. The major difference between this and previous learning algorithms is that we focus on "short" games that are repeated 100 rounds or less. Rather than developing an algorithm that has guaranteed long run convergence properties, we develop an algorithm that can generate data that resembles the data of human subjects in the experimental lab for these "short" games.

The task of developing computational agents could shed new light onto the theory of learning in games. The theory literature on learning in games has typically focused on learning algorithms with guaranteed convergence to some set, or small superset, of the Nash equilibria. The task of creating agents that mimic humans is different. By thinking of this new task, we may gain new insight into the decision making process in repeated games. For example, the algorithm proposed in this paper uses a pattern recognition scheme. Most of the theoretical models assume that players do not recognize patterns, even though patterns have been observed in the experimental laboratory. Focusing on this new task may help uncover aspects that were previously missing in the theory.

This computational model of learning can also complement experimental economics. We use computational agents to run simulations for both of these coordination games for a variety of experimental parameters. From these simulations, we develop hypotheses on human behavior which are then tested in the lab. We can run the simulations without any hypothesis in mind, and then develop these testable hypotheses from the simulation data. This alternative approach provides new insights into the human decision making process that have not been examined in the theoretical literature.

The rest of the paper proceeds as follows. Section 2.2 discusses the related work in this area. Section 2.3 lays out the model. This includes the setup of the games and equilibria, as well as a detailed description of the learning algorithm. Section 2.4 compares previous experimental data to simulation data from our algorithm as well as other learning algorithms. It also includes simulation results for the minimum-effort coordination game which are used to develop testable hypotheses for experiments. In section 2.5 these testable hypotheses are tested in the experimental laboratory. This section gives the experimental design as well as a comparison of the experimental and simulation results. We provide concluding remarks in Section 2.6.

2.2 Related Work

The theoretical literature on learning in repeated games started with the introduction of fictitious play by Brown (1951). In each round of fictitious play, each player best responds to the empirical distribution of past plays. Fictitious play has been shown to converge to Nash equilibria beliefs for two player zero-sum games (Brown 1951), 2 x 2 games (Miyasawa 1961), potential games (Monderer and Shapley 1996), and many classes of supermodular games (Milgrom and Roberts 1991, Hahn 2008, Berger 2007). Fictitious play has garnered much interest from game theorists because of its simplicity and its desirable convergence properties. However, fictitious play does not have desirable equilibrium selection properties, because once the players reach an equilibrium, they remain in that equilibrium forever.

Another common approach is to use evolutionary game theory and examine the evolutionary stable strategies (ESS) as introduced by Smith and Price (1973). In these models, strategies with random mutations compete against each other based on a fitness measure. Using an evolutionary model, Friedman (1991) shows that ESS are a subset of Nash equilibria. Adding randomness to the evolutionary model, Kandori, Mailath, and Rob (1993) and Young (1993) show convergence to risk dominant equilibrium in 2 x 2 symmetric games. Hart (2002) introduces a model in which the unique ESS leads to the subgame perfect Nash equilibrium. These evolutionary models provide stronger equilibrium selection properties than fictitious play, but they typically take thousands of rounds to attain these equilibria.

Most of the literature examining interaction in repeated games has focused on long run convergence properties; a few papers develop computational agents which act like human subjects.

Arifovic and Ledyard (2005) build computational agents to be used as a testbed for experiments on the Groves-Ledyard mechanism. In particular, the mechanism has one parameter which plays an important role in the speed of converge. They make predictions about optimal values of this parameter with their computational testbed, and then confirm these predictions with experiments. Their learning algorithm is a combination of a genetic algorithm with some behavioral intuition. Their computational agents are able to converge quickly, on average 20 rounds. However, their algorithm strongly favors convergence to a every agent playing the same action repeatedly. Our algorithm is able to get quick convergence in some games, but in addition it also converges to more complex strategies in other games.

One of the key features of our algorithm is a pattern recognition scheme. Pattern recognition is an important aspect of human decision process. Sonsino and Sirota (2003) find that over half of the human subjects converge to patterns of Nash equilibria (where different NE are played in subsequent rounds) in a variation of the Battle of the Sexes game. The only way that these patterns of Nash equilibria can be sustained is if players are recognizing patterns. However, with the exception of a few papers, previous learning models have not allowed agents to recognize these patterns in repeated interactions.

Sonsino (1997) proposes a theoretical model in which players recognize patterns, and play converges to a sequence of Nash equilibria with probability one. This would include the alternations observed experimentally in the Battle of the Sexes. However, he does not address how these patterns arise, rather that the players are able to recognize them once they have been played. In another theoretical paper, Lambson and Probst (2004) create a model where players find similar patterns in the history, and best respond to these patterns. Their learning algorithm converges to the convex hull of the set of Nash equilibria in games where fictitious play converges to the set of Nash equilibria. This paper does not give a specific technique for determining the length of patterns that players recognize. Rather, they assume that players only recognize patterns of a fixed length, and then examine the convergence of these players. Unlike these theoretical pattern recognition models, we actually implement computational agents which have a specific mechanism to recognize patterns of different lengths and select the optimal pattern lengths. Therefore, these agents are able to converge to single Nash equilibria as well as more complex patterns of Nash equilibria.

2.3 Model

2.3.1 Game

Before introducing the algorithm, we introduce the general game. The algorithm works on any variation of this general game. The game consists of n players, $N = \{1, 2, \ldots, n\}$. The set of strategy profiles is $S = S_1 \times S_2 \times \cdots \times S_n$, where player i's action comes from the infinite strategy space $s_i \in S_i = [0, 1]$. A specific strategy profile is the set of all players' actions, denoted by s. Let s_{-i} denote the actions of all players other than player i in strategy profile s. Each player has a payoff function which maps strategy profiles to real numbers, $\pi_i : S \to \mathbb{R}$.

The analysis in this paper focuses on two specific games from the general class of games described above: the minimum-effort coordination game and the Battle of the Sexes game. Both of these games are generalized to the continuous action space. The minimum-effort coordination game has n players, $N = \{1, \ldots, n\}$, each of which exerts a costly effort $s_i \in S_i = [0, 1]$. The payoff for player i is a term proportional to the minimum effort of all players, minus the cost of their own efforts. More formally,

$$\pi_i \left(s_i, s_{-i}\right) = \min_{j \in N} s_j - c s_i,$$

for some cost parameter c. The continuous version of the Battle of the Sexes game has two players, $N = \{1, 2\}$. Each player chooses a location on the unit interval, $s_i \in [0, 1]$. Based on the players' locations, the payoff function is,

$$\pi_1 \left(s_1, s_2\right) = \left(2s_1 - 1 - b\right)\left(2s_2 - 1 - b\right)$$

$$\pi_2 \left(s_1, s_2\right) = \left(2s_1 - 1 + b\right)\left(2s_2 - 1 + b\right),$$

for some parameter b. To get a better understanding of the Battle of the Sexes payoff functions,

	0	1
0	1,1	-1,-1
1	-1,-1	1,1

$b = 0$

	0	1
0	$9/4$, $1/4$	$-3/4$, $-3/4$
1	$-3/4$, $-3/4$	$1/4$, $9/4$

$b = .5$

	0	1
0	4,0	0,0
1	0,0	0,4

$b = 1$

Figure 2.1: Battle of the Sexes game for different values of b.

consider the 2-by-2 game where the payoffs are the same, but the actions space is now restricted to two actions, $S_i = \{0, 1\}$. The payoff matrices for different values of b are shown in Figure 2.1. There are a few things to notice about variations in b. First, it is symmetric around 0, i.e. the game is the same for $b = 0.1$ and $b = -0.1$ except that the players roles are switched. Without loss of generality, we only consider $b \geq 0$. When b is close to zero, the payoff difference between the equilibria at $(0, 0)$ and $(1, 1)$ is small while the cost of not coordinating (playing $(0, 1)$ or $(1, 0)$) is high. When b is close to 1, then the payoff difference between the two equilibria is large, but the cost of not coordinating is low. So this b parameter can be thought of as a symmetry parameter.

2.3.2 Equilibria

2.3.2.1 Minimum-Effort Coordination Game

The set of Nash equilibria of the minimum-effort coordination game depends on the parameter c. If $c > 1$, then the cost of effort is greater than the benefit, so it is a dominant strategy for each player to play the minimum effort, $s_i = 0$. If $c < 0$, the cost of effort is negative, so it is a dominant strategy for each player to exert the maximum effort, $s_i = 1$. The case that is usually studied is $0 \leq c \leq 1$. In this case, any strategy where all players exert the same effort is a Nash equilibrium.

The set of Nash equilibria is therefore defined by,

$$NE = \{s|s_i = s_j \text{for all } i, j \in N\}.$$

Since the strategy space is infinite, there are an infinite number of Nash equilibria. These equilibria are Pareto ranked. The equilibrium where all players exert the lowest effort is the Pareto-worst equilibrium, as each player gets payoff $\pi_i(0, 0, \ldots, 0) = 0$. The Pareto-optimal equilibrium is the equilibrium in which each player exerts the maximum effort and gets payoff $\pi_i(1, 1, \ldots, 1) = 1 - c$.

A player can always guarantee a payoff 0 by playing 0, so the minmax payoff in this game is 0 for all players. By the folk theorem for repeated games, this implies that any strategy that guarantees the players at least 0 is an equilibrium in the infinitely repeated game.

2.3.2.2 Battle of the Sexes Game

In the continuous Battle of the Sexes game, when $b > 1$, each player has a strictly dominant strategy, so there is a unique equilibrium. The more interesting case occurs when $0 \leq b \leq 1$, in which there are three equilibria. Given that player 2 is located at s_2, then the best response function for player 1 is as follows,

$$BR_1(s_2) = \begin{cases} 0 & s_2 < \frac{1+b}{2} \\ [0, 1] & s_2 = \frac{1+b}{2} \\ 1 & s_2 > \frac{1+b}{2} \end{cases}.$$

Given that player 1 is located at s_1, then the best response function for player 2 is,

$$BR_2(s_1) = \begin{cases} 0 & s_1 \leq \frac{1-b}{2} \\ [0, 1] & s_1 = \frac{1-b}{2} \\ 1 & s_1 > \frac{1-b}{2} \end{cases}.$$

Therefore, the three Nash equilibria are $(s_1, s_2) = (0, 0), \left(\frac{1-b}{2}, \frac{1+b}{2}\right), (1, 1)$. The equilibrium $(s_1, s_2) = \left(\frac{1-b}{2}, \frac{1+b}{2}\right)$ is Pareto-dominated by the other two equilibria. However, the other two equilibria are not Pareto rankable. Player 1 prefers the equilibrium $(s_1, s_2) = (0, 0)$ and player 2 prefers the equilibrium $(s_1, s_2) = (1, 1)$.

In this game, each player has a strategy which guarantees a payoff of 0, and this turns out to be the minmax payoff. By playing $s_1 = \frac{1-b}{2}$, player 1 will get a payoff of zero regardless of what player 2 plays. Similarly, player 2 can guarantee zero by playing $s_2 = \frac{1+b}{2}$. By the folk theorem, this implies that any strategy where the players are guaranteed a positive average payoff is an equilibrium of the infinitely repeated game. This includes the strategy where the players alternate between the two endpoints, which gives an average payoff of $1/2 \left[(1-b)^2 + (1+b)^2\right] = 1 + b^2$ per round.

2.3.3 Algorithm

In each period of a repeated game, the algorithm determines which action each agent plays. This action depends on the history of play as well as the current state of the agent. After each agent has made their choice, the actions and payoffs are revealed to all agents. The agents then update their history and current state, and choose their action the following round.

Two main features of this algorithm are the pattern recognition scheme and the agent's states. The experiments of Sonsino and Sirota (2003) show that human subjects are able to sustain non-trivial patterns of Nash equilibria. Even in 2-by-2 games, the probability of sustaining a pattern of Nash equilibria for n rounds by random choice decreases exponentially as n increases, yet the subjects are still able to sustain these patterns. The human subjects' ability to sustain these patterns of equilibria provide evidence that they are in fact recognizing these patterns. Therefore, pattern recognition is a natural feature when modeling human behavior in repeated interactions. Our pattern recognition scheme is a modification of the k-nearest neighbor classification algorithm from machine learning (Dasarathy 1991). Patterns are recognized by first identifying the current play (the most recent actions in the history) and then finding previous plays that are similar to the current play. The prediction for the next round is a weighted average of the outcomes of these similar plays. In each round, the agent makes a choice based on the current state, which is given by two parameters, γ and σ. The γ parameter represents the current level of confidence for an agent. This is determined by how well that agent is predicting what the other agents will do. The σ parameter represents the agent's satisfaction of the current play of the game. If the agent is not satisfied, and wants to change what is happening in the game, then σ is close to 1. If the agent is satisfied with how the game is going then σ is close to 0. When all agents have high values of γ and low values of σ, then each agent's choice has low variance and each agent is satisfied with the predicted outcome of his choice, so the algorithm has converged.

Another important aspect in the algorithm is that the agents are not able to calculate exact best-responses to their predictions. Instead, the agents determine best responses by randomly sampling from the strategy space, and keeping the strategy that gives the highest payoff. This is important for two reasons. First, it allows for completely general payoff functions. Since the explicit best response function is not required, the payoff functions need not be continuous or differentiable. Also, it allows agents to have different levels of intelligence by changing the number of samples they take. For example, an intelligent agent has a good grasp of the payoff function, and therefore is able to find the best response. This can be modeled by an agent who takes a large number of random samples to find the best response. Conversely, an unintelligent agent is not able to find the best response. This can be modeled as an agent that takes a small number of samples to find the best response.

The explanation of the algorithm is divided up into four parts: notation, preliminary initialization, round k action, and preparation for round $k+1$. For notational purposes, the superscript typically denotes the agent and the subscript denotes the round.

2.3.3.1 Notation

Each agent has a database of information that is used to help make their choice in each round. At the start of each round, each agent has two parameters in their database, the confidence parameter γ and the satisfaction parameter σ. These parameters for agent i in round k are denoted by γ_k^i and σ_k^i. The agents use these parameters to help make their choice. Agent i's choice in round k is represented by $\mathbf{x}_k(i)$. The choice of all agents in round k is given by $\mathbf{x}_k$, which yields payoffs

$\pi_i(\mathbf{x}_k) = \pi_k^i$ for agent i.

After making a choice, the agent updates his database of information in preparation for the next round. Each agent makes a prediction about what the other agents will play in the following round. Let $\hat{\mathbf{x}}_k^i(j)$ be agent i's prediction for agent j's play in round k. The full prediction vector, $\hat{\mathbf{x}}_k^i$, consists of predictions for all of the other agents.

As the game progresses, each agent creates a quasi-best-response matrix. Agent i's quasi-best response matrix at round k is denoted by Q_k^i. This matrix helps the agent determine which action to choose based on his prediction. To do this, the agent groups similar strategy profiles together in the quasi-best-response matrix. The agent then determines which play is best against these similar strategy profiles by randomly sampling responses from the strategy space. In the future, when a similar strategy profile arises, the agent uses this quasi-best-response matrix to help remember what he did in the past. From this quasi-best-response matrix, the agent determines the quasi-best-response for his prediction for round k, which is denoted by ${x_k^i}^*$. More details about the quasi-best-response are given below in the description of the algorithm in the preparation for round $k+1$ section.

Each agent also keeps track of his best and worst outcomes. To do this, each agent randomly chooses J strategy profiles from the uniform distribution on the joint strategy space $S = [0,1]^N$. Next, the agent calculates the payoffs for each of these profiles, and saves the strategy profiles which yield the highest and lowest payoffs, $\bar{\mathbf{x}}_k^i$ and $\underline{\mathbf{x}}_k^i$ respectively. These are referred to as the highest and lowest known choices for agent i in round k. The payoffs for these strategy profiles, $\bar{\pi}_k^i$ and $\underline{\pi}_k^i$, are referred to as the highest and lowest known payoffs for agent i at round k.

All of this information is stored in a database, and is available when the agent is making his choice in round k.

2.3.3.2 Initialization

Many learning algorithms contain multiple initialization periods, where the agents choose randomly in the strategy space. Since the focus of this paper is not long run convergence, but rather short run behavior, the initialization period has to be short. Before the first choice is made, the agents randomly choose J strategy profiles to determine their initial highest and lowest known payoffs, $\bar{\mathbf{x}}_0^i$ and $\underline{\mathbf{x}}_0^i$ respectively. Each agent then makes the initial predictions about the other agents by randomly drawing a number from the uniform distribution on $[0,1]$, that is $\mathbf{x}_1^i(j) \sim U[0,1]$. Finally, each agent starts with the lowest possible confidence level, $\gamma_1^i = 10$. The agents also start with the highest satisfaction parameter, $\sigma_1^i = 0$, because they have no reason to try to change the outcome of the game yet. With these initial parameters, the algorithm is ready to run.

2.3.3.3 Round k

Entering round k, agent i has a database of information which is used to make a choice in round k. The choice in round k is a random number drawn from a beta distribution with mean μ and variance ν^2. The mean of the distribution is a convex combination of the quasi-best-response, ${x^i}^*_k$, and the strategy which yields highest known payoff for agent i at round k, $\bar{x}_k^i$. The weight on each term is determined by the current level of satisfaction. If the agent's satisfaction level is high ($\sigma_k^i = 1$) then he plays the quasi-best-response for his prediction. If the agent is not satisfied ($\sigma_k^i < 1$), then he tries to move the outcome towards the point which yields his highest known payoff. That is,

$$\mu = \sigma_i^k {x^i}^*_k + (1 - \sigma_i^k)\bar{x}_k^i.$$

The variance of the distribution is inversely proportional to the current level of confidence[1]. The proportionality constant is ρ, so the variance is,

$$\nu^2 = \frac{1}{\rho \gamma_k^i}.$$

As the confidence level of the agent increases, the choice distribution has lower variance, and therefore the choice is more accurate. When the agent is not confident about what the other agents will do,

[1] It is not possible to have a distribution over a closed region, if the variance is high, and the the mean is sufficiently close to the endpoints. If this is the case, then it is corrected by using a modified beta distribution with mass point on the endpoint.

then the choice distribution has high variance, and the choice is not as accurate.

After all agents have made their choices as described above, the payoffs are calculated. The agents then learn the choices of the other agents as well as the payoffs of all agents. At this point, the agents begin their preparation for round $k+1$ by updating their database of information.

2.3.3.4 Preparation for Round $k+1$

The agents have a variety of tasks to perform in preparation for round $k+1$.

Update extremes As the game progresses, the agents become more acquainted with the payoff function. To model this, each round the agents update their highest and lowest known payoffs by taking J random samples from the joint strategy space. For each random sample $\mathbf{z}_j$, the payoff vector is calculated. If the payoff for agent i from the sample is higher than the highest known payoff for agent i in round k, then the agent sets the highest known choice for round $k+1$ to $\bar{\mathbf{x}}_{k+1}^i = \mathbf{z}_j$ and the highest know payoff round $k+1$ to $\bar{\pi}_{k+1}^i = \pi_i(\mathbf{z}_j)$. If none of the payoffs from the J sample points are higher than the highest known payoff for agent i at round k, then the highest known choice and payoff from round k are carried over to round $k+1$, i.e. $\bar{\mathbf{x}}_{k+1}^i = \bar{\mathbf{x}}_k^i$ and $\bar{\pi}_{k+1}^i = \bar{\pi}_k^i$. The same update is performed for the lowest known play and payoff.

Prediction for round $k+1$ In order to make a choice in round $k+1$, the agents make some prediction about what their opponents are going to do in round $k+1$. The prediction scheme used by the agents is a modification of the nearest neighbor classification algorithm from machine learning. The goal of the prediction scheme is to make a prediction for $\mathbf{x}_{k+1}$. Since there are N agents, the agents choices at round k are given by the vector $\mathbf{x}_k \in \mathbb{R}^N$. A pattern is vector combining one or more of these choice vectors. For example, a pattern of length 3 is $[\mathbf{x}_{k-2}\ \ \mathbf{x}_{k-1}\ \ \mathbf{x}_k]$. The agents divide the history of choice into the current pattern, previous patterns, and outcomes. Each previous pattern has a corresponding outcome. The algorithm makes a prediction for the outcome of the current pattern. The agents determine which of the previous patterns are closest to the current pattern. Then, the agents prediction is a weighted sum of the outcomes of the closest patterns. The agents repeat this process for patterns of different lengths, n. After this has been done for all values of n, the agent compares them, and determines which pattern length provides the best prediction.

For example, consider a two player game with the history of play after eight rounds,

Round	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Play	$(0,0)$	$(1,1)$	$(1,1)$	$(0,0)$	$(1,1)$	$(1,1)$	$(0,0)$	$(1,1)$	?

Consider the prediction by agent 1 of what agent 2 will play in the ninth round. First, agent 1 considers patterns of length 1. The current pattern is the most recent play, $(1,1)$. This has been played four previous times in rounds 2,3,5 and 6. These are the closest patterns. When these closest patterns have been played in the past, agent 2 has responded by playing 1,0,1, and 0 in the respective following rounds. These are the outcomes for the four closest patterns. This is not good, because agent 2 has played 0 half the time, and 1 half the time, so it is difficult to predict what agent 2 will play in the next round based on patterns of length 1.

Next, agent 1 looks at patterns of length 2. The current pattern in this case is the play in the previous 2 rounds, $(0,0),(1,1)$. This pattern has been played twice before in the past, in rounds 1-2 and 4-5. In response to this pattern, agent 2 has played 1 in both rounds 3 and 6. After patterns of length 2, agent 2 always chose 1. Therefore, patterns of length 2 are better for prediction that patterns of length 1.

More formally, at the end of the k^{th} round, each agent considers patterns of different lengths n. For each length, there are $k-n$ previous patterns of length nN each. The agent forms the previous patterns matrix $X \in \mathbb{R}^{k-n \times nN}$ and the output matrix $Y \in \mathbb{R}^{k-n \times N}$,

$$X = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ \mathbf{x}_2 & \cdots & \mathbf{x}_{n+1} \\ \vdots & & \vdots \\ \mathbf{x}_{k-n} & \cdots & \mathbf{x}_{k-1} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_2 \end{bmatrix} \text{ and } Y = \begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{x}_{n+2} \\ \vdots \\ \mathbf{x}_k \end{bmatrix}.$$

Each row of the previous patterns matrix is a single pattern, and these are denoted by X_m for $m = 1, \ldots, k-n$. The current prediction is the vector,

$$c = \begin{bmatrix} \mathbf{x}_{k-n+1} & \cdots & \mathbf{x}_k \end{bmatrix} \in \mathbb{R}^{nN}.$$

Next, the agent finds the j rows of X which are closest to the current pattern c in terms of Euclidean distance. To do this, the agent forms the distance vector by finding the length between the current point and each of the previous points,

$$\mathbf{d} = \begin{bmatrix} \|X_1 - c\| \\ \|X_2 - c\| \\ \vdots \\ \|X_{k-n} - c\| \end{bmatrix}.$$

Let L be the set of indices of the l smallest terms in the distance vector $\mathbf{d}$. That is $d_i \leq d_k$ for $i \in L$ and $k \notin L$. These indices correspond to the l rows of X which are closest to the current point c.

The agent now determines which pattern length gives the best prediction. As exhibited in the above example, the agent wants to choose the pattern length with the most similar outcomes. To determine the optimal pattern length for each n, the agent takes the outcome of the l closest points, and calculates the average of these points,

$$\bar{Y} = \frac{\sum_{l \in L} Y_l}{L}.$$

Then the agent computes the variance of these l closest points,

$$V_n = \sum_{l \in L} \|Y_l - \bar{Y}\|.$$

Now, the agent compares the variance for all considered pattern lengths and chooses the pattern length with the smallest variance. If there is a tie, then the agent chooses the shorter pattern. Note that average variances are higher in higher dimensions. This is not corrected for, which gives an additional benefit to the shorter patterns, because shorter patterns are easier to recognize.

Once the pattern length has been selected, the agent forms a weighted average of the closest outcomes. The closer the pattern is to the current outcome, the higher the weight is. The patterns are weighted using a logistic function. The prediction for the next period is thus,

$$\hat{\mathbf{x}}_{k+1}^{i}=\frac{\sum_{l\in L}Y_{l}e^{-\mathbf{d}(l)}}{\sum_{l\in L}e^{-\mathbf{d}(l)}}.$$

Therefore, if the distance to each of the l closest patterns is 0, then the prediction is just the average outcome from those l closest patterns. The agent makes his choice for period $k+1$ based on this prediction.

Quasi-Best-Response The quasi-best-response helps the agent determine the best response for his prediction for round $k+1$. To do this, the agent updates the quasi-best-response matrix from the previous period, Q_k^i. Each row of the quasi-best-response matrix consists of three items: prediction about what the other agents will do, what agent i should do given that prediction, and the payoff given that strategy profile. More formally row m has the terms,

$$Q_k^i=\left[\begin{array}{ccc}\mathbf{q}_{-i}^{m} & \mathbf{q}_{i}^{m} & \pi_i\left(\mathbf{q}_i^m,\mathbf{q}_{-i}^m\right)\end{array}\right].$$

Here, $\mathbf{q}_{-i}^{m}$ are the choices of the other agents, and $\mathbf{q}_{i}^{m}$ is the choice of agent i. Agent i updates Q_k^i as follows. First, agent i determines if the current prediction is similar to any of the entries already in the quasi-best-response matrix. To do this, agent i chooses a set, R, of random strategies. For each row of the quasi-best-response matrix, agent i calculates the payoff difference,

$$pd_m=\sum_{r\in R}\left|\pi_i\left(r,\mathbf{q}_{-i}^{m}\right)-\pi_i\left(r,\hat{\mathbf{x}}_{k+1}^{i}\right)\right|.$$

Next, the agents find the minimum payoff distance, $pd^*=\min pd_m$. If the distance is small, i.e. $pd^*<\delta$, then the two strategies are similar, and therefore are combined in the quasi-best-response matrix. If $pd^* > \delta$, then the two strategies are not similar, so a new entry is created in the quasi-best-response matrix. Let the threshold δ be a fraction of the difference between the highest and lowest payoff,

$$\delta = \frac{\bar{\pi}_i^k - \underline{\pi}_i^k}{20}.$$

If $pd^* < \delta$, then the agent updates the row of the quasi-best-response matrix corresponding to pd^*, call this row m^*. The agent takes the set of R strategies, and calculates the payoffs $\pi_i\left(r, \hat{\mathbf{x}}_i^{k+1}\right)$. Let r^* denote the strategy from R which maximizes $\pi_i\left(r, \hat{\mathbf{x}}_i^{k+1}\right)$. If this new strategy yields a higher payoff than the current quasi-best-response, i.e. $\pi_i\left(r^*, \hat{\mathbf{x}}_i^{k+1}\right) > \pi_i\left(\mathbf{q}_i^m, \mathbf{q}_{-i}^m\right)$, then the row m^* is updated to $\mathbf{q}_{-i}^{m^*} = \hat{\mathbf{x}}_i^{k+1}$ and $\mathbf{q}_i^{m^*} = r^*$.

If $pd^* > \delta$, then the agent creates a new row for the quasi-best-response matrix, call this row $M + 1$. Again, the agent calculates the payoffs $\pi_i\left(r, \hat{\mathbf{x}}_i^{k+1}\right)$ for all $r \in R$, with r^* being the strategy which yields the maximum payoff. The agent then updates the quasi-best-response matrix by setting $\mathbf{q}_{-i}^{M+1} = \hat{\mathbf{x}}_i^{k+1}$ and $\mathbf{q}_i^{M+1} = r^*$.

Update γ The parameter γ measures the current level of confidence of the agent. When accurate predictions are made, the agent's confidence increases. In preparation for round $k + 1$, the agent compares the prediction for round k that was made in round $k - 1$, $\hat{\mathbf{x}}_k^i$, with the actual play from round k, $\mathbf{x}_k$. Based on this prediction and outcome, the agent updates his confidence as follows,

$$\gamma_{k+1} = \frac{\alpha_1}{\|\hat{\mathbf{x}}_k^i - \mathbf{x}_k\| + \alpha_2}\gamma_k.$$

Therefore, if the Euclidean distance between the prediction and the actual outcome is less than $\alpha_1 - \alpha_2$, then his confidence increases. The maximum possible increase in confidence is α_1/α_2.

Update σ The parameter σ represents the agent's satisfaction at the current state of the game. If not satisfied with the current outcome, the agent may try to induce the other agents to play something else in order to change the current outcome. If this attempt to move is unsuccessful, the agent will stop trying. For example, suppose two agents are coordinating at one of the equilibria repeatedly in the Battle of the Sexes game. Agent 1 is at her optimal equilibrium, and Agent 2 is at his least favored equilibrium. Agent 2 realizes that he can receive a higher payoff at the other equilibrium. Therefore he will try to induce Agent 1 to start playing the other equilibrium. However, Agent 1 may not change the way she is playing, even when Agent 2 starts playing something else. If Agent 2 has tried for a long time with no success, he will give up, and start playing the original equilibrium. The entire process of trying to move and giving up is called a *moving session*.

Agent i starts with the highest satisfaction possible. The satisfaction remains at the highest level until some event causes agent i to start a moving session. To become dissatisfied, the agent has to have a good idea of what the other agents are going to play. Therefore, agent i must have a confidence greater than γ_{MS} in order to start a moving session. Given that confidence is greater than γ_{MS}, agent i starts a moving session in two situations. If agent i knows that all agents receive higher payoffs at his highest known play, then he tries to move there because everyone will receive a higher payoff. Also, if agent i's highest known payoff increases his payoff by a large amount, and decreases the other agents' payoffs by only a small amount, then he tries to change the outcome. There are also some situations in which agent i does not start a moving session when his confidence is greater than γ_{MS}. If moving to agent i's highest known play increases agent i's payoff by a small amount, but decreases all other agents payoffs by a large amount, then agent i does not try to change the outcome. Also, if agent i has tried to move before unsuccessfully, then he will not try to move again until he has found a better strategy.

Once the moving session has started, agent i tries to induce the other agents to play his optimal strategy. If the play of the game is moving away from the play at the start of the moving session, and towards the highest play for agent i, then agent i will continue the moving session. If the play of the game does not move towards the highest play for agent i, then that round will be considered a *failure*. If the total number of failures becomes to high, then i will stop the moving session.

To more formally define this event that triggers a moving session, consider the term,

$$\Sigma_k^i = \frac{\pi_i\left(\bar{\mathbf{x}}_k^i\right) - \pi_i\left(\mathbf{x}_k\right)}{\frac{1}{N-1}\sum_{j\neq i}\pi_j\left(\bar{\mathbf{x}}_k^i\right) - \pi_j\left(\mathbf{x}_k\right)}.$$

Σ_k^i is referred to as the *relative gain* for agent i in round k. Agent i's payoff at the highest known play is always greater than his payoff at the current play, because the agent takes the current play into account when updating his highest known play. Therefore, switching from the current play $\mathbf{x}_k$ to agent i's highest known play $\hat{\mathbf{x}}_k^i$ will always increase agent i's payoff. So the numerator of Σ_k^i is always positive.

The agent also keeps track of the *maximum relative gain* for round k, $\bar{\Sigma}_k^i$, and the *minimum relative gain* for round k, $\underline{\Sigma}_k^i$. At the beginning of the game, agent i starts with maximum relative gain of $\bar{\Sigma}_0^i = 0$ and minimum relative gain of $\underline{\Sigma}_0^i = -1$. The agent updates these extreme relative gains with the current relative gain when the current relative gain is more extreme (higher than maximum or lower than minimum) and confidence is greater than γ_{MS}. The role of the extreme relative gains is to ensure that the agent does not continuously try to move to a point which the other agents refuse to move to.

Based on the current relative gain, the extremes relative gains, and the confidence, agent i determines whether or not to start a moving session. When the denominator of Σ_k^i is positive, and hence $\Sigma_k^i > 0$, the other agents benefit on average when switching from $\mathbf{x}_k$ to $\bar{\mathbf{x}}_k^i$. So, if $\Sigma_k^i > \bar{\Sigma}_k^i$ and $\gamma_k^i > \gamma_{MS}$, then the agent starts a moving session because all agents will receive higher payoffs at $\bar{\mathbf{x}}_k^i$. When the denominator of Σ_k^i is negative, the other agents get lower payoffs on average when switching from $\mathbf{x}_k$ to $\bar{\mathbf{x}}_k^i$. However, if Σ_k^i is very negative, then the average decrease of the other agents payoff is small compared to the increase for agent i. So if $\Sigma_k^i < \underline{\Sigma}_k^i$ and $\gamma_k^i > \gamma_{MS}$ then the agent also starts a moving session. To summarize, agent i tries to move if $\Sigma_k^i \notin \left[\underline{\Sigma}_k^i, \bar{\Sigma}_k^i\right]$ and $\gamma_k^i > \gamma_{MS}$.

In the first round of the moving session, agent i decreases from the full satisfaction level $\sigma = 1$ to the level $\sigma = \sigma_0 < 1$. Agent i also sets the number of failures to 0, $f = 0$. Agent i should not expect the other agents to respond to this move until they have seen the play in second round of the moving session and had at chance to respond to it in the third round of the moving session. So the agent remains with satisfaction $\sigma = \sigma_0$ in the second round of the moving session, and this does not count as a failure. Starting in the third round, agent i's satisfaction and failures depend on whether the other agents are responding to agent i's move. In particular, if the other agents are responding, and play is moving toward the highest known payoff, i.e.,

$$\|\mathbf{x}_k - \bar{\mathbf{x}}_k^i\| > \|\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1}^i\|,$$

then the satisfaction increases, $\sigma_{k+1} = \bar{\xi}\sigma_k$ and the number of failures stays constant $f_{k+1} = f_k$ (for some $\bar{\xi} > 1$). Alternatively, if the other agents are not responding, so play is not moving toward agent i's highest know payoff, i.e.

$$\|\mathbf{x}_k - \bar{\mathbf{x}}_k^i\| < \|\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1}^i\|,$$

then the satisfaction decreases, $\sigma_{k+1} = \underline{\xi}\sigma_k$ and the number of failures increases by one, $f_{k+1} = f_k + 1$.

When the amount of failures reaches the threshold $f_k = \bar{f}$, then the session ends because the other agents are not responding to the move. After the session ends, the amount of failures is set back to 0, and the satisfaction is set back to the highest level $\sigma = 1$.

2.4 Simulation Results

To test the algorithm, we first compare simulations to some experimental results reported in the literature. We then compare the simulations of our algorithm to simulations using fictitious play and the algorithm proposed in Arifovic and Ledyard (2005). From now, fictitious play will be referred to as FP, Arifovic and Ledyard's (2005) algorithm as AL, and our algorithm as PR (pattern recognition).

2.4.1 Comparisons with Experiments

2.4.1.1 Minimum-Effort Coordination Game

There have been several papers that run experiments with minimum-effort coordination games with continuous strategy spaces. Goeree and Holt (2005) (GH from now on) run experiments on the

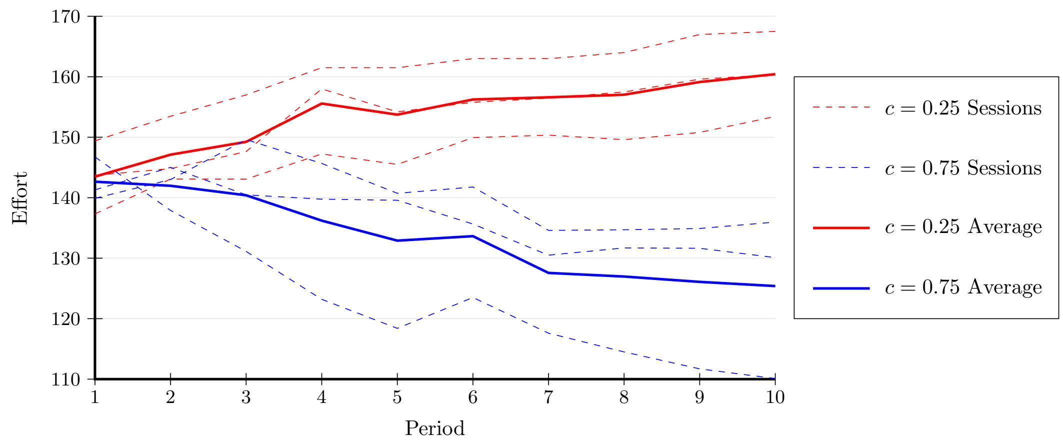


Figure 2.2: Experimental results of continuous minimum-effort coordination game from Goeree & Holt (2005).

continuous minimum-effort coordination game with group size $N = 2$. Each of their six sessions involved ten subjects who played a game for ten rounds. At the beginning of each round, each subject was randomly paired with another subject. Subjects were then asked to pick an effort level from the range $[110, 170]$. After all subjects had picked their effort, their payoffs and their partner's choice were revealed. They ran two treatments of three sessions apiece; one with low cost ($c = 0.25$), and another with high cost ($c = 0.75$). Their results are displayed in Figure 2.2. These experiments show that the effort levels of the subjects in the high cost treatments become progressively lower, and the effort levels of the subjects in the low cost treatments become progressively higher. In addition, the authors point out that late in the experiment, the choices in different treatments are separated by the mean of the range (140). GH also run similar treatments for 20 rounds, and find that the adjustment of choices tends to happen in the first 10 rounds, and the play levels out after that. These are the characteristics that are desired in the simulations using the PR algorithm.

The minimum-effort coordination game experiments run in GH can be simulated with computational agents using PR to make their decision. One difference between the experiment and our setup is that GH use different random matched pairs each round rather than fixed pairs throughout the experiment. To correct for this, we set the satisfaction parameter $\sigma = 0$. The satisfaction parameter allows the agents to induce the other agents to play some strategy. This parameter is set to zero

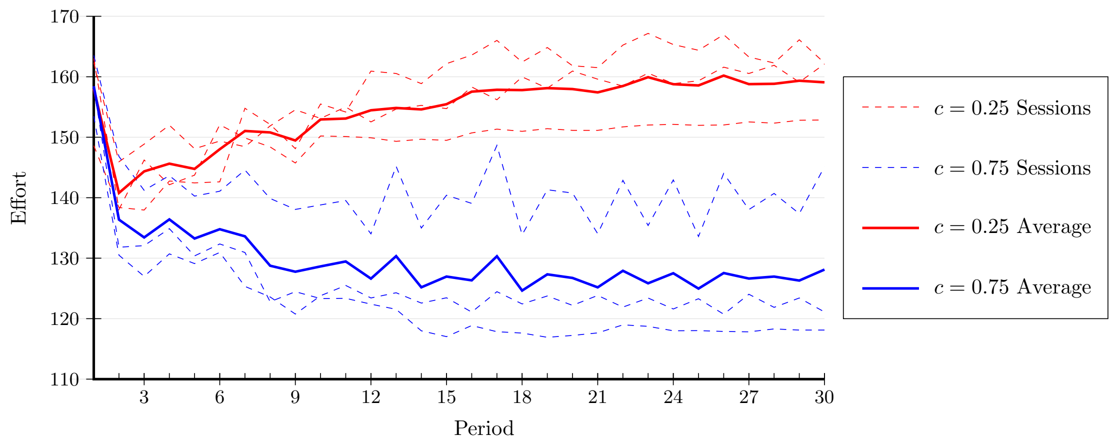


Figure 2.3: Simulation of continuous minimum-effort coordination game experiments with agents using PR.

because it is difficult for one agent to induce another agent to play something when they will both be randomly matched with another agent in the following round. The results from the simulations are presented in Figure 2.3. As in the actual experiment, there is separation between the high and the low cost treatments. In addition, the separation of the two treatments occurs around the midpoint as observed in the experiments. Finally, most of the adjustment happens in the first 10 rounds, and play levels out after that. So the agents in the simulations exhibit similar dynamics and equilibrium selection properties to the human subjects in the continuous minimum-effort coordination game.

There have also been several papers which examine the effect of group size on coordination in the minimum-effort coordination game. Van Huyck, Battalio, and Beil (1990) (VHBB) run experiments with minimum-effort coordination games while varying group size[2]. They find that for large group sizes (N=14-16) with $c = 0.5$, subjects converge to the Pareto dominated equilibrium, where all players exert the lowest effort. They also find that after failing to coordinate on the Pareto superior equilibrium in large groups, these same subjects are able to converge to the Pareto superior equilibrium 12 out of 14 times when matched in fixed groups of two. So even when the subjects are used to playing the dominated equilibrium, they are still able to coordinate on the high effort when matched in fixed pairs. This provides strong evidence that fixed groups of two should be able to reach the Pareto superior equilibrium a large percentage of the time.

[2]Their subjects have 7 choices rather than a continuum of choices.

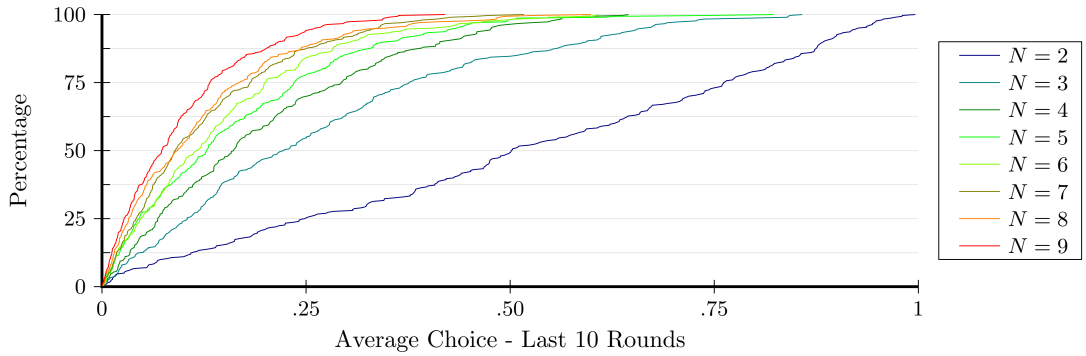
Knez and Camerer (1994) (CK) extend the result from VHBB by running similar experiments with more group sizes, and again using $c = 0.5$. They find that as the number of subjects increases, coordination on high effort strategies becomes more difficult. They also find that the biggest difference occurs when increasing the number of subjects from two subjects to three subjects. They provide two reasons for this. First, adding an additional player that chooses randomly leads to a weakly lower minimum. Second, the belief structure becomes ambiguous when increasing from two to three subjects, which causes players to be more cautious when choosing their effort level.

We run simulations of the minimum-effort coordination game with different group sizes using FP, AL, and PR to see if they match the experimental results above. For each group size, we run 300 simulations of 100 rounds each with cost $c = 1/2$. We show that the simulations using PR are more similar to the experimental data than the simulations using FP and AL. The details of these simulations are left to the appendix.

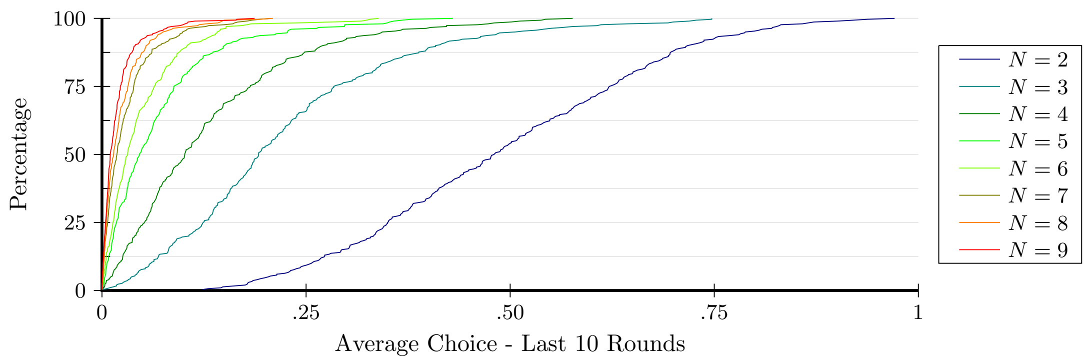
To examine the outcomes of these simulations, we look at the average choice in the last 10 rounds. By the last 10 rounds, every agent was playing within 0.05 of the other agents over 99% of the time with FP, over 98% of the time with AL, and over 99% of the time with PR. So the average over the last 10 periods is a reliable measure of the convergence of strategies[3]. This average choice in the last 10 rounds is referred to as the convergence point of the simulation. The empirical cumulative distribution functions of convergence points for each of these learning algorithms are displayed in Figure 2.4. More coordination on the high effort outcomes is represented by ECDfs that are further to the Southeast. The ECDF corresponding to perfect coordination on the high effort equilibrium would be 0% everywhere except 1. If the agents always converge on the low effort equilibrium, then the ECDF would be 100% everywhere except 0. The ECDFs show that the comparative statics on N match the experimental data for all three sets of simulations: as the number of agents increases, coordination on Pareto superior equilibria becomes more difficult.

The average convergence point for each group size is summarized in Figure 2.5. The average convergence point is the expected outcome for a simulation, and corresponds to the mean of the

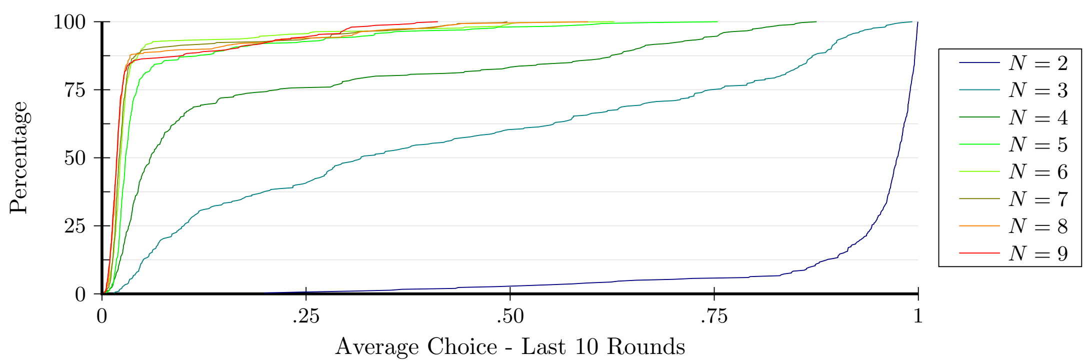
[3] The reason for using the average over the last 10 periods rather than the choice in the last period is to avoid problems where all of the agents have played the same thing for 9 rounds in a row, but in the last round, one agent changes actions.



(a) FP



(b) AL



(c) PR

Figure 2.4: Empirical cumulative distribution functions of convergence points for FP, AL, and PR for different group sizes.

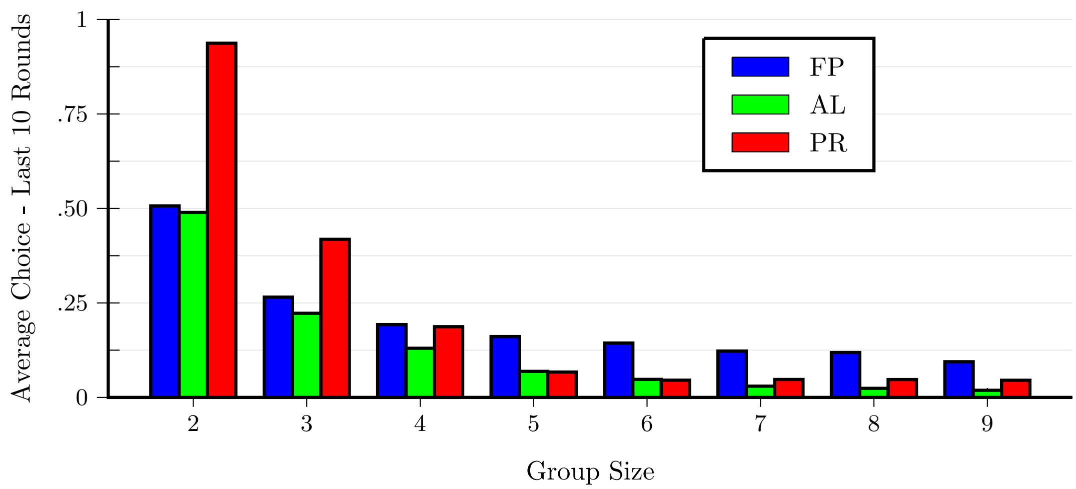


Figure 2.5: Average Convergence Points for FP, AL, and PR for different group sizes.

ECDF for that simulation. As shown in CK, the figure shows that in all three simulations, the biggest difference in ability to attain the high effort outcome occurs when increasing group size from two to three. The FP and AL simulations have relatively low convergence points for groups of size two. The average convergence point for simulations with groups of two for FP is 0.51 and for AL is 0.49. This is in contrast to the experiments run in VHBB which show that fixed pairs of agents are able to converge to the Pareto superior equilibrium about 85% of the time. The average convergence point of the PR simulations is 0.94 which is more in line with the coordination behavior exhibited in VHBB by human subjects in the laboratory. Based on Figure 2.5, the PR algorithm matches the experimental data better than both AL and FP.

2.4.1.2 Battle of the Sexes Game

Next we run simulations of the Battle of the Sexes game using the three different algorithms. The Battle of the Sexes game, like the minimum-effort coordination game, has multiple equilibria. But, in the Battle of the Sexes game, the equilibria are not Pareto ranked. One desirable outcome that has been observed in the experimental literature is when players are able to alternate between the two equilibria. This maintains the maximum welfare, but also yields equal payoffs for both players. While most of the literature has focused on more complex extensions of the Battle of the Sexes game, this outcome has been observed in the simple Battle of the Sexes game as in Rapoport, Guyer, and Gordon (1976) and more recently in Sonsino and Sirota (2003). To test to see if the learning algorithms can obtain this outcome, we run 300 simulations of 100 rounds each of the Battle of the Sexes game for each learning algorithm. Each game consists of two players, and has cost $c = 0.5$.

When agents use FP to make their choice, two outcomes occur. Either the players converge to one of the equilibria, or they start a cycle of non-coordination. In the non-coordination cycle, the players will always choose the opposite location, and therefore it is not a Nash equilibrium and the payoff is the lowest possible. Out of the 300 simulations, they reach this non-coordination cycle 41 times. Every other time they converge to one of the two equilibria. When agents use AL to make their choice, they have a tendency to converge to playing the same choice repeatedly. Therefore, it difficult to converge to patterns of equilibria. In the simulations with AL, agents converge to one of the two equilibria every time out of 300 trials. So neither FP or AL converge to the pattern of Nash equilibria that has been observed experimentally.

Finally we run simulations where the agents use PR to make their choices. To determine if agents converge to a pattern of equilibria, we develop two types of convergence. First, there is convergence in strategies. Convergence in strategies occurs at the round when the difference between any agent's choice in two consecutive rounds is less than 0.05 for the remainder of the game. This means the agents are playing the same thing repeatedly. The other type of convergence is convergence in γ, which happens when $\gamma > 5000$ for all agents for the remainder of the game. This means that all agents are making accurate predictions about what the other agents will do. Convergence to a pattern of Nash equilibria will occur when the agents converge in γ but not in strategies. Out of the 300 simulations when agents use PR to make their choice, they converge to a pattern of Nash equilibria 15 times. Out of these 15 patterns 14 consist of players playing one equilibrium in even rounds, and the other equilibrium in odd rounds. In the other pattern, agents repeatedly play one of the equilibria twice and the other equilibrium once. So agents that use PR are able to converge to patterns of Nash equilibria that have been observed by human subjects in experiments.

Our main focus was to develop a learning algorithm that could sustain these patterns of Nash equilibria, which PR is able to do. While these alternations have appeared in the experimental data, there are no experiments on the continuous version of the Battle of the Sexes game. Also the experiments on the discrete version typically have other aspects such as communication. Therefore, it is difficult to compare the results of these simulations with the current literature to any level beyond the presence of these patterns.

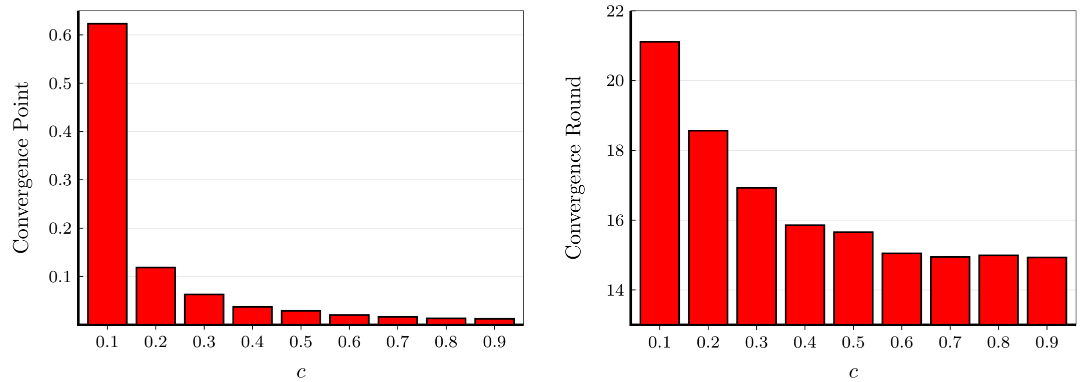
2.4.2 Experimental Hypotheses

Next, we run simulations using the algorithm on the minimum-effort coordination game and develop testable experimental hypotheses. The benefit of using computational agents is that simulations are essentially costless, which allows us to run many trials for a wide range of parameter values.

Previous experiments on the minimum-effort coordination game have focused on differences in cost and group size. The experiments have typically compared two different parameter values: a low and high cost or a small and large group (Goeree and Holt 2005). Experiments examining a large set of parameters are difficult due to constraints on the number of subjects in a given subject pool, as well as monetary costs for running large experiments. Simulations using the algorithm provide a testbed to simulate these experiments for many different parameter values. Unlike the binary comparisons, examining a larger set of parameters will give us a better understanding of the behavior which may have been overlooked in the past.

From the minimum-effort coordination game, we run simulations with groups of four agents with 9 different costs, varying from $c = 0.1$ to $c = 0.9$. At each parameter value, we run 300 simulations lasting for 50 rounds.

Convergence Point: We find that higher costs lead to lower convergence points. Convergence points are the average play over the last 10 periods of the repeated game. The convergence points of these simulations are displayed in Figure 2.6(a). This is consistent with experimental results from minimum-effort coordination games as shown in Goeree and Holt (2005).



(a) Convergence Point

(b) Convergence Round

Figure 2.6: Convergence Points and Convergence Round as a Function of Cost.

Convergence Speed: We then examine the effect of different costs on speed of convergence.[4] Based on the simulations, we find that the number of rounds required to converge decreases with c, so convergence is faster when the cost is higher. A plot of convergence as a function of c is displayed in Figure 2.6(b) (higher bars mean slower convergence). The intuition for increase in speed of convergence for higher cost is simple; it is more expensive for agents to search for different outcomes or experiment with different strategies.

Average Payoff: These convergence results have some interesting effects on the agents' payoffs. When agents do not all choose the same effort (i.e., best respond), the outcome is Pareto inefficient. If all agents chose the minimum effort for a given strategy profile, then everyone's payoff would be weakly higher, with at least one receiving a strictly higher payoff. Since it is inefficient when all agents are not choosing the same effort, slow convergence may lead to lower average payoffs. The average payoff per agent for different costs is displayed in Figure 2.7. It is difficult to compare the welfare between two experiments with different costs because they have different payoff functions. Even though welfare is difficult to compare, the payoff for any given strategy profile is lower when the cost of effort is higher. Intuition thus suggests that higher cost of effort should lead to lower average payoffs in the repeated game. However, the simulations suggest that higher cost may actually lead to higher payoffs.

[4]We use convergence in γ as a measure of convergence.

On one hand, with higher costs the agents receive lower payoffs for similar strategy profiles. In addition, the agents are converging to lower effort, which also leads to lower payoffs. However, the simulations show that high costs yield faster convergence, which eliminates some of the inefficiency caused by non-coordination, and increases the agents' payoffs. In fact, the increase in payoffs due to faster convergence outweighs the decrease in payoffs due to higher cost and lower convergence point. Note that the difference in average payoff shrinks as number of rounds increases in Figure 2.7. This result is due to the fact that the inefficiencies of non-convergence at the beginning of the repeated interaction become less important as the number of rounds increases, since it is more likely that the agents have converged.

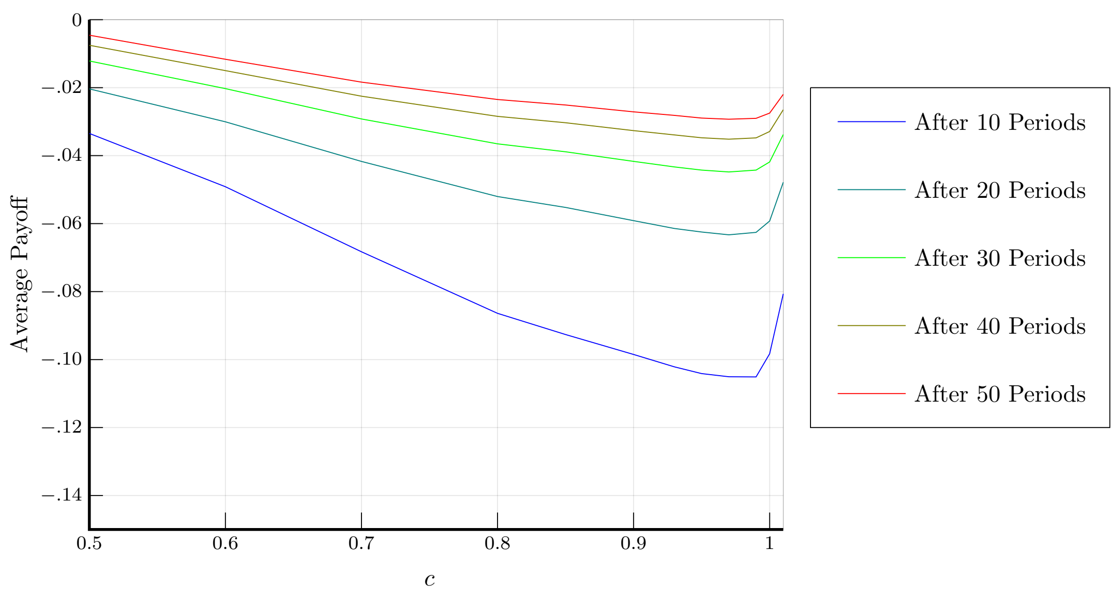


Figure 2.7: Average Payoffs for Different Costs in Minimum-Effort Coordination Game as a Function of Number of Rounds.

Based on the simulations, we test the following three hypotheses in the experimental laboratory:

Hypothesis 1. Convergence Point: The game will converge to lower effort levels as the cost of effort increases.

Hypothesis 2. Convergence Speed: The game will converge faster to an equilibrium as the cost of effort increases.

Hypothesis 3. Average Payoff: The average payoff does not monotonically decrease as the cost of effort increases.

2.5 Experiments

In this section we describe the setup and results for the experiments that will test the hypotheses formulated from the simulations.

2.5.1 Design

2.5.1.1 Overview

The experiments were conducted at the California Social Science Experimental Laboratory (CASSEL) located in the University of California, Los Angeles (UCLA). A total of 60 subjects participated in the experiments. The average performance-based payment was 20USD. All students were registered as subjects with CASSEL (signed a general consent form) and the experiment was approved by the local research ethics committee at both universities. These labs consist of over 30 working computers divided into a cubicles, which prevents students from viewing another student's screen.

The experiment was programmed and conducted with the experiment software z-Tree (Fischbacher 2007). The instructions were available both in print as well as on screen for the participants, and the experimenter explained the instruction in detail out-loud. Participants were also given a brief quiz after instruction to insure proper understanding of the game and the software. A copy of the instruction, as well as the payoff tables, are available in the Appendix.

The subjects were randomly assigned to their roles in the experiment. Furthermore, no one participated in more than one experiment. The identity of the participants as well as their individual decisions were kept as private information. However, each groups knew their own minimum effort. Experiment used fictitious currency called francs. The participants were fully aware of the sequence, payoff structure, and the length of the experiment. All participants filled out a survey immediately after the experiment.

2.5.1.2 Details of the Experiment

A total of 20 subjects participated in each session. These 20 subjects were split into 5 groups of 4, and each group used a different cost parameter. The entire session was divided into 5 blocks, and each block was divided into 15 rounds. After each block, the subjects were randomly rematched (with replacement) to another group of 4 and were randomly reassigned another payoff parameter (with replacement). See Figure 2.8 for the time line.

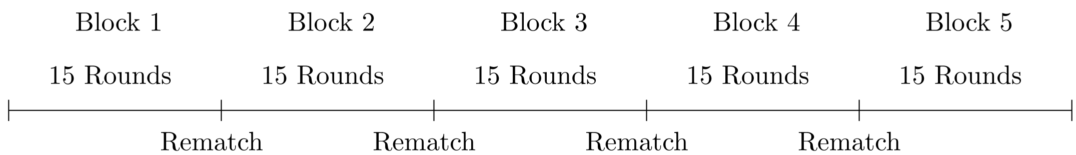


Figure 2.8: Timeline and Matching Structure for the Experiment.

Subjects played a minimum-effort coordination game each round. Their task was to choose an effort level, $s_i \in \{1, ..., 7\}$, and their payments were determined by the following payoff function

$$p_i = 1000 \left(\min_{j \in N} \{s_j\} \right) - c(s_i) + 5950$$

In each block, there were 5 groups each with a different payoff matrix based on $c \in \{50, 500, 900, 950, 990\}$. The subjects were shown the payoff table displayed in Table 2.1, with the calculation already completed for the subjects. The group size, randomization, and the fact that everyone in the group was using the same payoff table were common knowledge. However, the group's own minimum effort was private information to the group and was not available to the outside members.

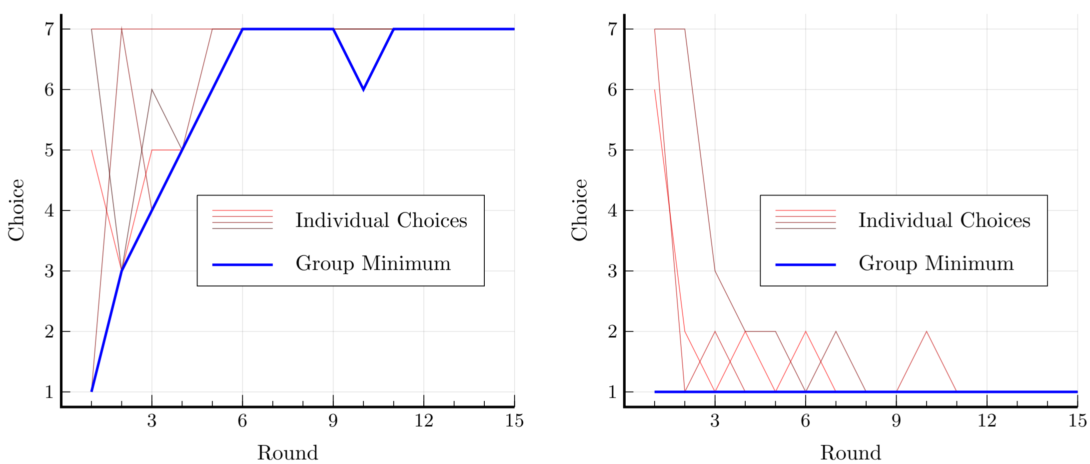
2.5.2 Results

Figure 2.9 illustrates sample results from one of the block of sessions. Figure 2.9 (a) is an example where the group coordinates on the high-effort equilibrium (converging to an effort level of 7) and Figure 2.9 (b) is an example where the group coordinates on the low-effort equilibrium (converging to an effort level of 1).

	Minimum effort of all agents						
***i*'s Effort**	7	6	5	4	3	2	1
7	$12950 - 7c$	$11950 - 7c$	$10950 - 7c$	$9950 - 7c$	$8950 - 7c$	$7950 - 7c$	$6950 - 7c$
6	–	$11950 - 6c$	$10950 - 6c$	$9950 - 6c$	$8950 - 6c$	$7950 - 6c$	$6950 - 6c$
5	–	–	$10950 - 5c$	$9950 - 5c$	$8950 - 5c$	$7950 - 5c$	$6950 - 5c$
4	–	–	–	$9950 - 4c$	$8950 - 4c$	$7950 - 4c$	$6950 - 4c$
3	–	–	–	–	$8950 - 3c$	$7950 - 3c$	$6950 - 3c$
2	–	–	–	–	–	$7950 - 2c$	$6950 - 2c$
1	–	–	–	–	–	–	$6950 - c$

Table 2.1: Sample Payoff Table that was used in the Experiment.

Calculations were already filled in for the subjects



(a) Coordination to High Effort (b) Coordination to Low Effort

Figure 2.9: Sample Results From One of The Blocks for Illustration Purpose.
The thin lines represent individual choices and the thick line represents the group's minimum choice

2.5.2.1 Convergence Points

First, we test the hypothesis that higher costs lead to lower convergence points and provide the results in Figure 2.10 (details in Table A.1 and Table A.2 in appendix). These results are taken from the average choice of the last 5 rounds and it supports the hypothesis that the average choice drops as the cost parameter increases. While the cost parameter $c \in \{50, 500\}$ provides a high level of average choice around 4.5 to 5, the average choice drops significantly lower to 1 to 1.2 for cost parameter $c \in \{900, 950, 990\}$. Although there is not a significant difference between the means from $c = 900$ and $c = 950$, the differences are significant in the right direction for the rest of the mean comparisons.

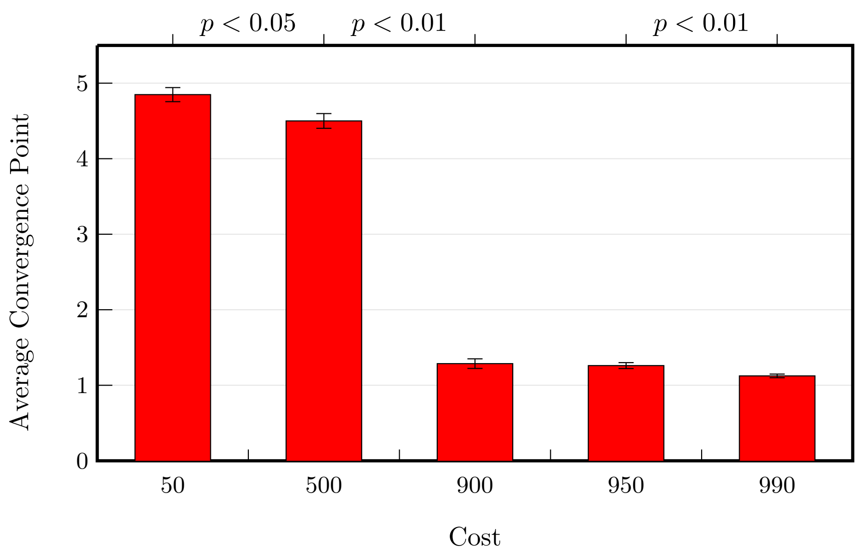


Figure 2.10: Average Convergence Points for Different Cost Parameters (See Tables A.1 and A.2 for numbers).

2.5.2.2 Convergence Speed

Comparing convergence speed is bit trickier than comparing convergence points. In the simulations, the agents are computer programs, so it is easy to see why they are making certain choices. However, with humans it is not as straightforward. Consider the following example in Figure 2.11. If one were to use a rule that the convergence occurs when there are no deviations (i.e., everyone is best responding), then there will not be any convergence until round 13 in the example. When studying experimental results with subjects from a laboratory, this may be too conservative of a criterion. Noisy choice in human behavior is often expected in experiments. Whether these noises are rational or not is another story. However, there are many different ways of modeling noisy choices, such as the Quantal Response Equilibrium (McKelvey and Palfrey 1995), the Level-K Model, and the Cognitive Hierarchy Model (Camerer, Ho, and Chong 2004), among others.

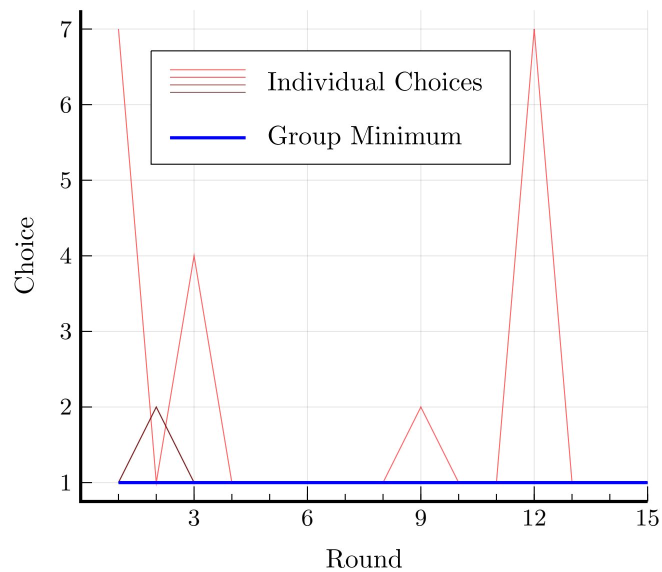


Figure 2.11: An example of the difficulty of measuring convergence.
The thin lines represent individual choices and the thick line represents the group's minimum choice

Here, we provide two means of measuring convergence. First, we use a more quantitative measure of convergence called v-bounded condition. Then we introduce a more qualitative and intuitive measure of convergence called the similarness condition.

Definition 2.5.1. *The game has converged to a particular equilibrium at round t under v-bounded condition if the variance of efforts chosen is always less than v for every round starting from t. Specifically, $var_{t+m}(\sigma_1, ..., \sigma_n) \leq v$, $\forall m \geq 0$.*

For example, if the strategy profile σ consists of $[3, 3, 3, 4]$, this will require that a variance parameter of $v \geq 0.25$ will be needed to consider this strategy profile as converged under the v-bounded condition. See Table 2.2 for other samples of strategy profile and its required variance parameter for v-bounded condition.

Using the v-bounded condition criterion for the notion of convergence, Figure 2.12 illustrates

$\boldsymbol{\sigma}$	**Minimum** v
$[3, 3, 3, 4]$	.25
$[3, 3, 4, 4]$	.33
$[2, 3, 3, 4]$	.66
$[3, 3, 4, 4]$	.92
$[3, 3, 3, 5]$	1

Table 2.2: Samples of Strategy Profile and its Required v Parameter for v-bounded Condition.

the average rounds it took for the game to converge.[5] Although convergence speed seems to be increasing as the cost parameter increases, differences are not statistically significant. To illustrate why the v-bounded condition may not be a good criterion, consider the following example from Figure 2.11: the condition requires $v \geq 9$ in order to allow this particular example to be considered converged due to a large jump in choice of effort by one of the players in round 12. This does not take into account that the deviation is by one person for only one period. However, intuitively, one may think that this game has converged at round 4.

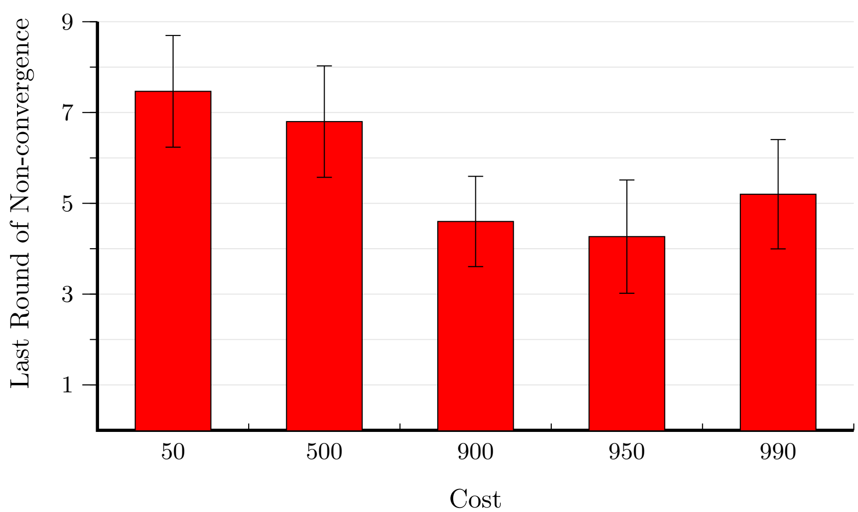


Figure 2.12: Number of Rounds Needed for Convergence for $v \leq 0.5$.

Therefore, we use a more intuitive and qualitative measure of convergence. For a given round, look at the number of different effort levels in the strategy profile. Then, the game has converged to a particular equilibrium if a high proportion of people use the same strategy. We call this the *similarness condition*. The added benefit of the similarness condition is that it does not unreasonably

[5]We drop the last round deviation because there may be end game effects.

penalize cases where one person may deviate significantly away from the best response for just one period. By the same token, it also means that this measure treats the following two strategy profiles as equally converged: $[2, 2, 2, 3]$ and $[1, 1, 1, 7]$.

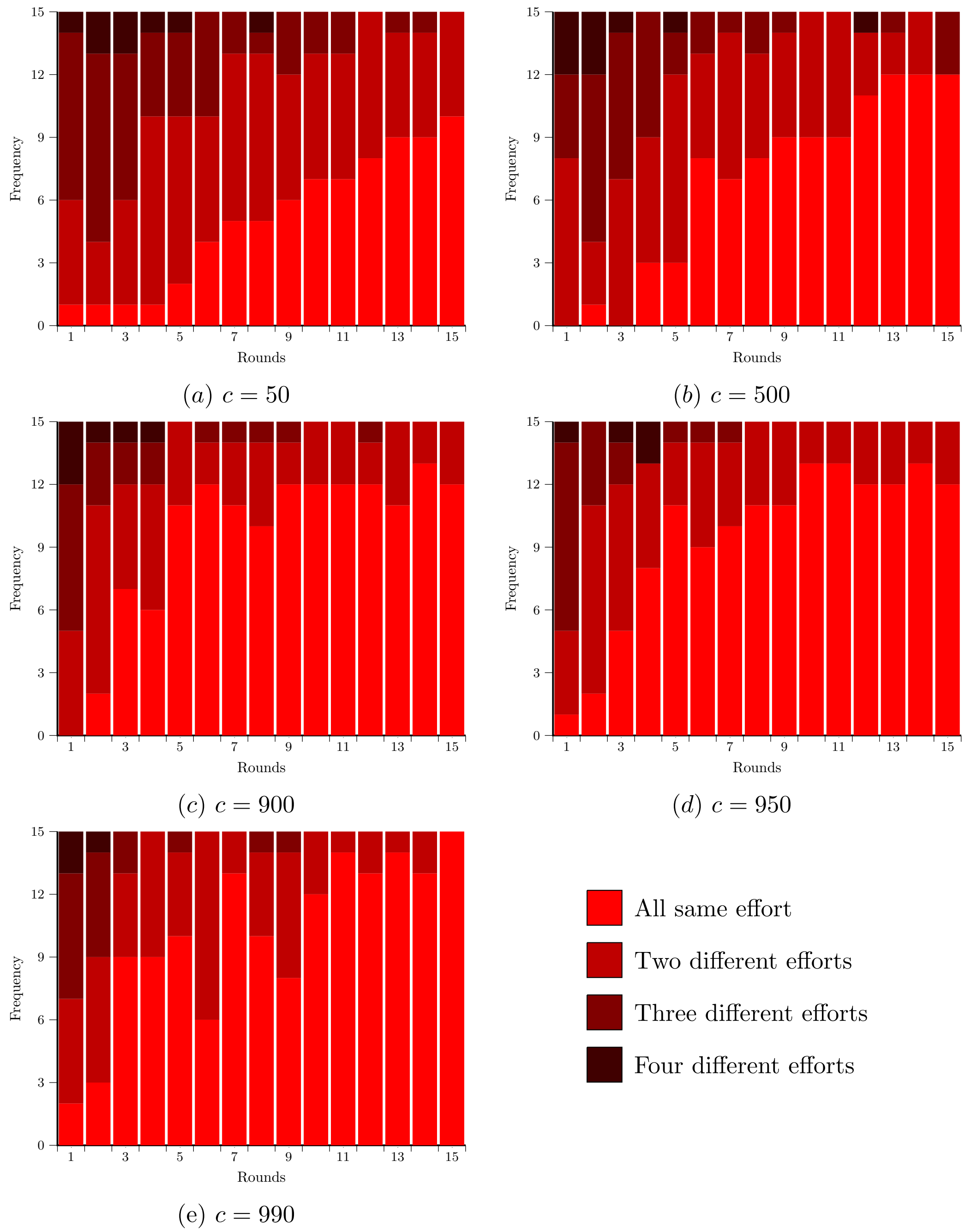
Figure 2.13 shows the frequency of different strategies played for various costs of effort. If the game is indeed converging faster under the similarness condition, we expect to see a higher frequency of all same effort (lightest) and two different efforts (light), which indicates everyone playing the same strategy and three people playing the same strategy, respectively. As the cost of effort increases, we observe an increase in frequency of all playing the same effort and only 2 different effort levels. This increase in frequency holds true for any given round. Furthermore, the frequency of one or two different effort levels also increases as the experiment proceeds (number of rounds increases). In other words, there are many different strategies being played in the initial round but subjects learn to best respond.

This similarness condition as a convergence criterion provides support that the game converges faster to a particular equilibrium as the cost of effort increases.

2.5.2.3 Average Payoff

Finally, we analyze the behavior of the average payoff as the cost increases. Refer to Figure 2.14 (exact numbers in Table A.3 and A.4 in appendix) to see the average payoff and their mean comparisons up to 4 rounds for each of the cost parameters from the experiment. We find a statistically significant decrease in average payoff from $\mu_{50} = 9088$ at $c = 50$ to $\mu_{950} = 4846$ at $c = 950$ ($p \approx 0$). However, as the simulation has predicted, the average payoff at $c = 990$ of $\mu_{990} = 5136$ is significantly higher than the average payoff at $c = 950$ of $\mu_{950} = 4846$ ($p < 0.05$). This suggests that early in the interaction the average payoffs are not monotonically decreasing as the costs increase.

Given that we observe a non-monotonicity in average payoff as a function of cost of effort in the first 4 rounds, we test the significance after the entire block of the experiment (15 rounds). The result is displayed in Figure 2.15 (see Tables A.5 and A.6 for exact numbers). Again, we observe a similar pattern to the results from the first 4 rounds. The average payoff of $\mu_{990} = 5650$ at $c = 990$



(a) $c = 50$ (b) $c = 500$ (c) $c = 900$ (d) $c = 950$ (e) $c = 990$

Figure 2.13: Frequency of Different Strategies Played for Various Costs.

is significantly greater than the average payoff of $\mu_{950} = 5560$ at $c = 950$ ($p < 0.1$). Furthermore, the average payoff in this setting is the lowest at $c = 950$, which is also lower than the average payoff of $\mu_{900} = 5652$ at $c = 900$ ($p < 0.1$).

While the p-values for the non-monotonicity hypothesis are weaker after 15 rounds than after 4 rounds, which is what the simulations predicted. The difference between the average payoff when $c = 990$ and $c = 950$ diminishes as more rounds are played. This confirms the prediction made by the simulation in Figure 2.7. As more rounds are played, the positive welfare from the lower cost averages out the negative welfare from the wasted effort. For example, after 4 rounds, the difference in average payoff is $\mu_{990} - \mu_{950} = 288.9583$. But, after 15 rounds, the difference decreases to $\mu_{990} - \mu_{950} = 90.1889$. In other words, the non-monotonicity of average payoff is most salient at the initial phase of the game.

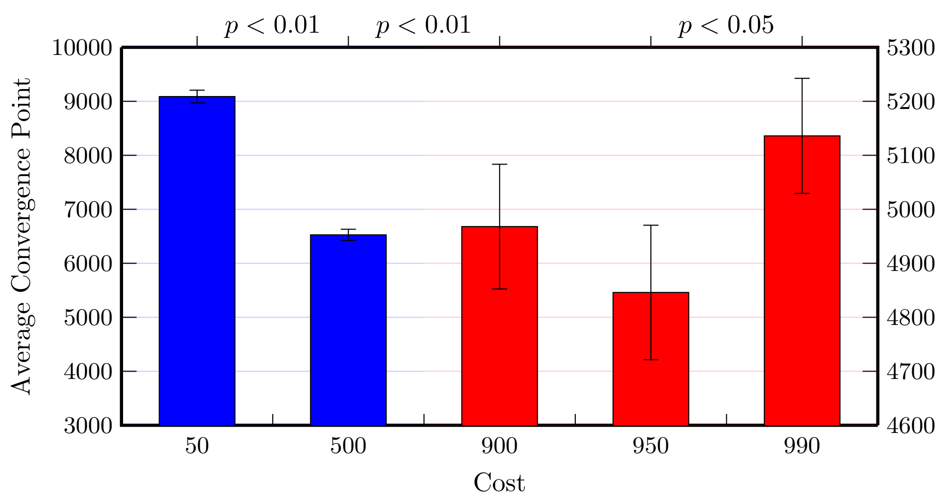


Figure 2.14: Average Payoff After 4 Rounds (See Tables A.3 and A.4 for numbers).

2.6 Conclusion

We have developed a learning algorithm which can be implemented with computational agents. We have also shown that simulations with these computational agents are able to generate data which shares many features with experimental data for both the minimum-effort coordination game and the Battle of the Sexes game. In particular, agents using our algorithm are able to sustain alternations

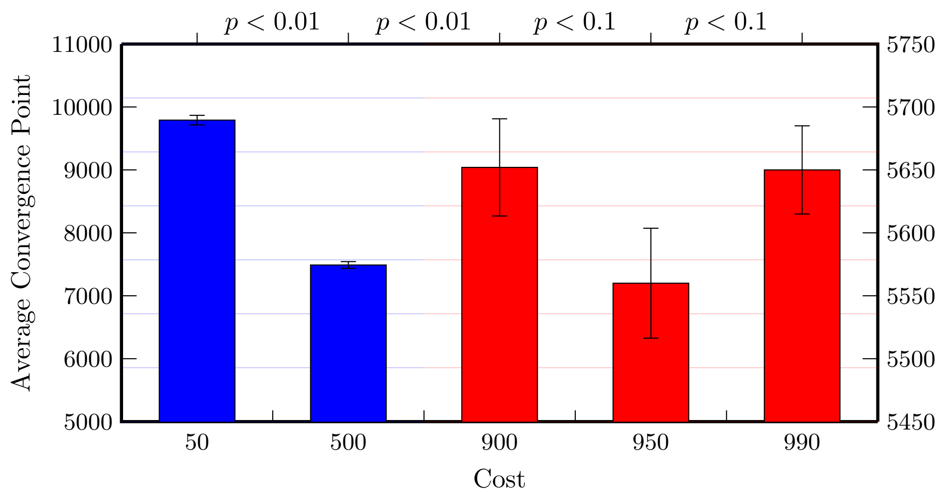


Figure 2.15: Average Payoff After 15 Rounds (See Tables A.3 and A.4 for numbers).

between Nash equilibria in the Battle of the Sexes game: a result which previous learning models have had a difficult time sustaining. We have used these agents to run simulations for a wide class of parameters in the minimum-effort coordination game, and developed testable hypotheses about human behavior in this game. We have designed an experiment based on these simulations and confirmed the testable hypotheses developed from the simulations.

The validity of our algorithm has been tested with experimental evidence from two coordination games. Ideally, this algorithm will provide results consistent with a much wider class of games. The wider the class of games that this type of algorithm is consistent with, the more reliable the results will be for simulations of new games. As these agents become a more reliable predictor of human behavior, they can become very valuable asset to experimental economists. In the beginning we can replace costly pilot experiments by running simulations of experiments with these agents to determine optimal parameter values for experiments. As these agents become more reliable, we may be able to use these agents to interact with human subjects during experiments, providing larger subject pools. Eventually, entire experiments could be run with these computational agents, and when a useful result is found, it can be confirmed experimentally with human subjects.

Chapter 3

Bounded Rationality in Repeated Games

3.1 Introduction

Models of bounded rationality assume that agents have limited ability to process information and solve complex problems (Simon 1957). These models are often able to make sharper predictions than their fully rational counterparts (Conlisk 1996). When players are fully rational and they interact repeatedly, a plethora of equilibrium outcomes are possible. In particular, these games suffer from folk theorems; namely any individually rational and feasible payoff[1] is attainable in equilibrium. This multitude of equilibria suggests further analysis of the equilibrium selection problem is needed. This paper focuses on the question: Does a model of repeated interactions with boundedly rational agents lead to a smaller set of outcomes in equilibrium? A smaller set of outcomes is required to make better predictions about what type of behavior we should expect to see in repeated interactions.

In this paper, players are limited in two ways. First, as in many repeated interactions, players are not able to see the actions of their opponents. Rather, they get an imperfect signal from which the action must be inferred. An example is the "secret price cutting" game (Stigler 1964), in which two competing firms give unobservable price cuts to their customers, which can only be inferred through sales figures. In this paper, each player receives a private signal that correctly conveys the action of their opponent with probability (accuracy) less than one. This builds on the literature

[1] A payoff is feasible if there is some infinite sequence of actions which leads to this payoff. A payoff is individually rational if each player is receiving at least their minmax payoff. The minmax payoff is the payoff that minimizes that maximum possible loss. See Mailath and Samuelson (2006) for more information.

that examines imperfect private monitoring in repeated games (Kandori 2002).

The second limitation involves memory constraints. Typical repeated game strategies require that players have perfect memory, and can differentiate between every possible infinitely repeated game history[2]. Due to memory constraints, it is inconceivable that any economic agent could differentiate between every history in this infinite set. Here, I assume that players are able to classify this infinite set of histories into a finite number of groups (referred to as states). This leads to an intuitive class of strategies that capture the simple heuristics used during the infinitely repeated game.

By limiting recall to finite states, I can represent players' strategies by finite automata. Intuitively, a finite automaton can be thought of as a set of mental states. Each state represents a different mood (for example good and bad), and therefore may lead to a different behavior (nice and mean). Based on the actions of the other player, the mood might change, in which case the automaton moves (transitions) to a different state. A more sophisticated player may have many states which represent a complex strategy, while a simple player may have only a handful of states. Representing strategies with finite automata was first suggested by Aumann (1981) and has been widely studied since (Mailath and Samuelson 2006).

Using agents that are limited in the ways described above, I examine a model of repeated games, where it is possible for players to attain cooperative relationships without using contracts. The main insight from this paper is that in equilibrium players must play strategies that attain cooperation, and are forgiving enough to avoid long conflicts if cooperation breaks down. Conflicts between players are suboptimal, so the time spent in these conflicts has to be short in relation to the time spent in cooperation. I show that if players spend long periods of time in conflict, then it is possible for them to switch their strategy to something that avoids conflict.

I first consider the case where players select automata with no more than two states. In this case, the set of equilibrium strategies is small. For a class of infinitely repeated prisoner's dilemma games, there are at most two types of equilibria when signal accuracy is less than one (Theorem 3.4.3). In the first type of equilibrium strategy, a fixed sequence of actions is played regardless of the action

[2]A history refers to a sequence of previously played actions.

of the opponent. The other type of equilibrium strategy follows the simple heuristic: if the other player cooperates, continue playing the same action; if the other player defects, switch actions. This simple strategy, introduced in the theoretical biology literature, has been coined "win-stay, lose-shift" (WSLS) (Nowak and Sigmund 1993). If both players play WSLS, then high levels of cooperation are attained. WSLS is special because it is forgiving, and allows for quick recoordination after cooperation breaks down. I also give sufficient conditions on stage-game payoffs which guarantee that both players playing WSLS is an equilibrium when signal accuracy is sufficiently close to one (Theorem 3.4.4). These sufficient conditions hold for a large class of 2×2 games, suggesting that WSLS is a useful strategy in a wide variety of settings. Finally, experiments run by Wedekind and Milinski (1996) using human subjects suggest that WSLS is played in repeated prisoner's dilemma games. When players are limited to two-state automata, the number of outcomes is small, and the predictions are supported by experimental evidence.

Next, I examine a more general model in which players' strategies may be any finite-state automaton. In this case, I give necessary and sufficient conditions for the structure of equilibrium strategies when signal accuracy is close to one (Theorems 3.5.6 and 3.5.8). These conditions formalize the insight that players must spend almost all the time cooperating. To prove these conditions, I show that if players are not cooperating most of the time, then there exist better strategies which allow players to avoid long periods of conflict and spend almost all the time cooperating. These results show that the benefits of WSLS (cooperation and recoordination) are still required in equilibrium in a more general model.

The analysis presented here is most similar to Compte and Postlewaite (2009). They examine an infinitely repeated prisoner's dilemma game where players have imperfect private monitoring and choose among two-state automata. They show that cooperation is possible for a large region of accuracy and payoff combinations. They focus on a specific class of strategies: two-state automata with fixed transitions. Players choose only the action to be played in each state. In addition they allow for a common knowledge public signal (generalized to almost public signal) which allows the players to coordinate their actions.

The results considered in this paper extend those in Compte and Postlewaite (2009) in several ways. First, the two-state results consider the entire set of two-state automata. This means players select their transition functions as well as an action to be played in each state. I find that WSLS is able to attain high levels of cooperation even when there is no public signal to aid recoordination. I also show that WSLS remains an equilibrium over a large class of 2×2 games. In addition, I examine a more general model which allows for any finite-state automata.

This paper proceeds as follows. In Section 3.2, I give a motivating example, which highlights the problems of imperfect monitoring. Then, in Section 3.3, I present the model of boundedly rational agents and define the equilibrium concept. Next I give the results of the paper. First, in Section 3.4, I consider the restricted case where players only choose among two-state automata. Then in Section 3.5, I consider the case where players can choose among any finite-state automata. In Section 3.6.1, I provide a brief review of some related literature. Finally, I conclude and provide extensions in Section 3.7.

3.2 Motivating Example

In 1980, Robert Axelrod invited a number of top scholars to submit programs to compete in an iterated Prisoner's Dilemma tournament. The strategy that fared best was tit-for-tat, which simply repeats the play of the opponent in the previous round (Axelrod 1980a, Axelrod 1980b). In later work, Axelrod suggested that players may not perfectly perceive their opponents actions. To further examine the effect of misperceptions, he ran simulations where players had a 1 percent chance of seeing the incorrect action of their opponent. Not surprisingly, he found that these misperceptions led to lower levels of cooperation. However, tit-for-tat was still the dominant strategy in the tournament. Axelrod notes,

> "[TIT FOR TAT] got into a lot of trouble when a single misunderstanding led to a long echo of alternating retaliations, it could often end the echo with another misperception. Many other rules were less forgiving, so that once they got into trouble, they less often got out of it. TIT FOR TAT did well in the face of misperception of the past because it could readily forgive and thereby have a chance to reestablish mutual cooperation."
>
> -Axelrod (2006)

	C	D
C	1,1	-L,1+L
D	1+L,-L	0,0

Figure 3.1: Game PD.

This excerpt captures one of the main insights of this paper: Players do not want to play strategies which get stuck in suboptimal periods. Axelrod states that tit-for-tat was successful because it was forgiving enough to be able to avoid these suboptimal periods more than most strategies. However, the following example shows that these suboptimal periods can be detrimental to payoffs, even when both players play tit-for-tat.

Consider two players playing the infinitely repeated prisoner's dilemma game displayed in Figure 3.1. Each player plays the tit-for-tat strategy. Each player starts by cooperating, and then repeats their opponents play from the previous round. If players' signals are perfect, they continue to play C throughout the remainder of the repeated game. Based on the payoff table, this leads to an average payoff of 1 per round.

Now, suppose that players receive an imperfect signal about their opponents action. The players start by cooperating. They continue to cooperate as long as the signals are correct. Suppose that in round r, both players play C, but player 1 gets an incorrect signal that player 2 played D. In round $r+1$ player 1 plays D because of the incorrect signal, and player 2 continues to play C. If both players receive correct signals in round $r+2$, then player 1 plays C and player 2 plays D. The players continue to "echo" each other's action until another incorrect signal is received. While stuck in the period of alternations, the average payoff for each player is $1/2$, lower than the payoff when cooperating.

If during this period of alternations, one player receives a signal that C was played when actually D was played, then both players perceive the actions as C, and hence both cooperate in the following round. This cooperation continues until another incorrect signal is received. However, if one of the players receives a signal that D was played when actually C was played, both players perceive action D, and both will defect in the following round. This mutual defection continues until at least one player receives an incorrect signal. The average payoff per round when both players are defecting is 0.

When both players play tit-for-tat, there are three periods the system can get stuck in: always play C, echo alternations, or always play D. The only way to get out of one of these periods is if one of the players receives an erroneous signal. Suppose the signal is correct with probability $1-\varepsilon$ and incorrect with probability ε. Over the course of the infinitely repeated game, for all $\varepsilon > 0$, the frequency of time spent in the cooperate and defect periods is $1/4$ and the alternating period is $1/2$. Therefore, the frequency of each of the four possible action combinations is equal in the infinitely repeated game. So each player gets an average payoff of $1/4\left[u_i\left(C,C\right)+u_i\left(C,D\right),u_i\left(D,C\right)+u_i\left(D,D\right)\right] = \frac{1+(1+L)-L}{4} = \frac{1}{2}$ in every round. Both players would receive higher payoffs if they played cooperate all the time.

In Section 3.4 , I show that in contrast to tit-for-tat, when both players play “win-stay, lose-shift”, the system does not get caught in these suboptimal periods. When an incorrect signal is received, the strategies are able to recoordinate quickly without incurring large losses. Then, in Section 3.5, I show that in a more general model, players still do not play strategies that get stuck in suboptimal periods in equilibrium. Before the results, I first introduce the formal model and some notation.

3.3 Model

Two players, $\mathcal{I} = \{1,2\}$, play the supergame G. In every round, the players play the stage game $g = \{S_1, S_2, u_1, u_2\}$. In the stage game, each player has $|S_i|$ pure strategies. The stage-game payoff function is $u_i : S_1 \times S_2 \to \mathbb{R}$. The stage-game payoffs for player i can be represented by a payoff matrix $P_i \in \mathbb{R}^{|S_1|\times|S_2|}$. In the supergame G, the agents play stage game g for an infinite number of rounds $t = 1, 2, 3, \ldots$.

3.3.1 Imperfect Monitoring

After both players have their chosen actions in round t of the supergame, each player receives a private signal which conveys the true action of their opponent with probability less than one. More formally, with probability $r_i(s_1, s_2, \varepsilon)$ player i receives a signal that the other player played action s_2 when the other player actually played s_1. The signals functions have a common rate of error, $\varepsilon \in [0, .5]$. For example, if $S_1 = S_2 = \{C, D\}$, one possible signal function is

$$\begin{aligned} r_i(C, C, \varepsilon) = r_i(D, D, \varepsilon) = 1 - \varepsilon \\ r_i(C, D, \varepsilon) = r_i(D, C, \varepsilon) = \varepsilon. \end{aligned} \tag{3.1}$$

In words, the signal is correct with probability $1 - \varepsilon$ and incorrect with probability ε. This signal function is referred to as the simple signal function, r_i^S, and is used for many examples and results in this paper.

3.3.2 Imperfect Memory

Players have bounds on their ability to differentiate between infinitely repeated game histories. Players are only able to classify this infinite set of histories into a finite number of states. This restriction yields a simple set of strategies which can be represented with finite-state automata.

A finite automaton is defined as a quadruple, $M = \left(Q_i, q_i^0, f_i, \tau_i\right)$. Here, Q_i is the finite set of states for player i and q_i^0 is the initial state. In each state, the automaton prescribes a pure action, which is determined by the output function $f_i : Q_i \to S_i$. Finally, the transition function determines which state to transition to based on the current state and the action of the other player, $\tau_i : Q_i \times S_{-i} \to Q_i$. Since the output function depends on S_i and the transition function depends on S_{-i}, if players have different action sets, then each player selects from a different set of automata. The set of all finite automata for player i is denoted by $\mathcal{M}_i$.

At the beginning of the supergame, each player chooses a finite-state automaton. After each

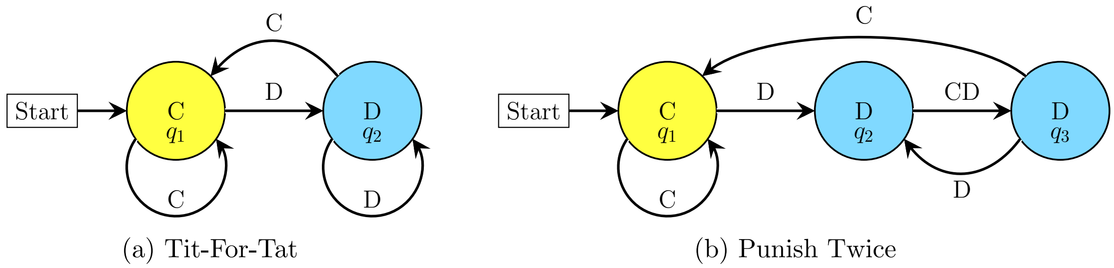


Figure 3.2: Examples of Automata.

history, this automaton is in a certain state, and plays the action corresponding to that state. So a finite automaton prescribes a stage-game action for every possible history.

Examples of Automata

Some example automata are displayed in Figure 3.2. The first automaton represents the tit-for-tat strategy. There are two states; the automaton cooperates in the first state, and defects in the second state. When the other player plays C, this automata goes to the first state, and hence this player cooperates. When the other player plays D, this automata leads to the second state, so this player defects. The second automaton represents a punish twice strategy. The first state of this automaton is a cooperative state. The automaton remains in this state as long as the other automaton is playing C. When the other automaton defects, this automaton goes into a two-state punishment phase. In the first state of the punishment phase, the automaton plays D and regardless of the other play goes to the third state. In the third state, the automaton plays D, and returns to the cooperative state only if the other automaton plays C. More complex strategies, such as N-period action sampling (Selten and Chmura 2008), can also be represented with a finite automaton.

3.3.3 Payoffs and Equilibria

When choosing automata, the players try to maximize the non-discounted limit of means. For a given pair of signal functions, the payoff is determined by the choice of automata from each player, and the level of error in signal function ε, $U_i : \mathcal{M}_1 \times \mathcal{M}_2 \times [0, .5] \to \mathbb{R}$. Given the signal function, error level, and automata, there is some infinite sequence of realized joint actions, $x^0, x^1, \ldots$. The players payoff is the average payoff per round over this infinite sequence of joint actions.

$$U_i\left(M_1, M_2, \varepsilon\right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} u_i\left(x^t\right),$$

where $u_i\left(x^t\right)$ is the payoff for player i when joint action x^t is played.

In this paper I assume non-discounted payoffs.[3] This allows me to focus on long run equilibrium rules-of-thumb rather than strategies where players deviate in the beginning because they are impatient.

Definition 3.3.1 (Best Response)**.** *Player i's best response function* $BR_i : \mathcal{M}_{-i} \times [0, .5] \to \mathcal{M}_i$

$$U_i\left(BR_i(M, \varepsilon), M, \varepsilon\right) \geq U_i\left(M', M, \varepsilon\right) \text{ for all } M' \in \mathcal{M}_i.$$

Definition 3.3.2 (Nash Equilibrium)**.** *For fixed signal functions* r_i *and error level* ε*, a pair of automata,* (M_1, M_2)*, is an equilibrium of the supergame G if and only if* $M_i = BR_i\left(M_{-i}, \varepsilon\right)$ *for* $i = 1, 2$.

A Nash equilibrium pair of automata is referred to as an equilibrium.

3.4 Two-State Automata

In this section, I analyze the set of equilibria when players strategies are restricted to two-state automata. First, I introduce some important automata. I then show that for a class of infinitely repeated prisoner's dilemma games, there are at most two types of equilibria for any parameter pair. I then give sufficient conditions on stage-game strategies that ensure that WSLS is an equilibrium for all small error levels. Finally, I discuss some previous work done on WSLS, including some experiments from Wedekind and Milinski (1996) which show that human subjects play these strategies in the laboratory.

[3] Here I assume that players payoff is determined by the limit of means. Limit of means can sometime be problematic because the limit may cease to exist in some cases, leading to an incomplete preference order. This however is not a problem here as displayed in Lemma 3.5.3.

3.4.1 Important Two-State Automata

The restricted set of automata, $\mathcal{M}^2$, consists of only two-state automata. For notational simplicity, automata are represented by a tuple with the starting points omitted,

$$M = (\{f(q_1), \ldots, f(q_n)\}, \{\tau(q_1, C), \ldots, \tau(q_n, C)\}, \{\tau(q_2, D), \ldots, \tau(q_n, D)\}).$$

The starting points are mentioned when relevant. Before giving a characterization of the two-state equilibria, I first need to introduce some automata.

- **Always play C** - $M^C = (\{C\}, \{q_1\}, \{q_1\})$
- **Always play D** - $M^D = (\{D\}, \{q_1\}, \{q_1\})$
- **Alternating** - $M^{CD} = (\{C, D\}, \{q_2, q_1\}, \{q_2, q_1\})$
- **Win-Stay, Lose-Shift** - $M^{WSLS} = (\{C, D\}, \{q_1, q_2\} \{q_2, q_1\})$

Automata "always play C" and "always play D" play the same action regardless of the signal. The alternating automaton always alternates between C and D regardless of the signal. The "win-stay, lose-shift" automaton follows the simple rule: if I get a signal that you cooperated, then I play the same action; if I get a signal that you defected, I switch actions. These automata are displayed in Figure 3.3.

It is important to point out the differences between M^{WSLS}, and the well-studied tit-for-tat automaton, $M^{TFT} = (\{C, D\}, \{q_1, q_1\}, \{q_2, q_2\})$. Both automata yield output C in state q_1 and D in state q_2. The transitions from the first state of these two automata are the same as well. Both remain in the state if the signal is C, and change if the signal is D. The difference occurs in the second state transitions. The second state of M^{TFT} returns with a D and changes with a C while the second state of M^{WSLS} returns with a C and switches with a D.

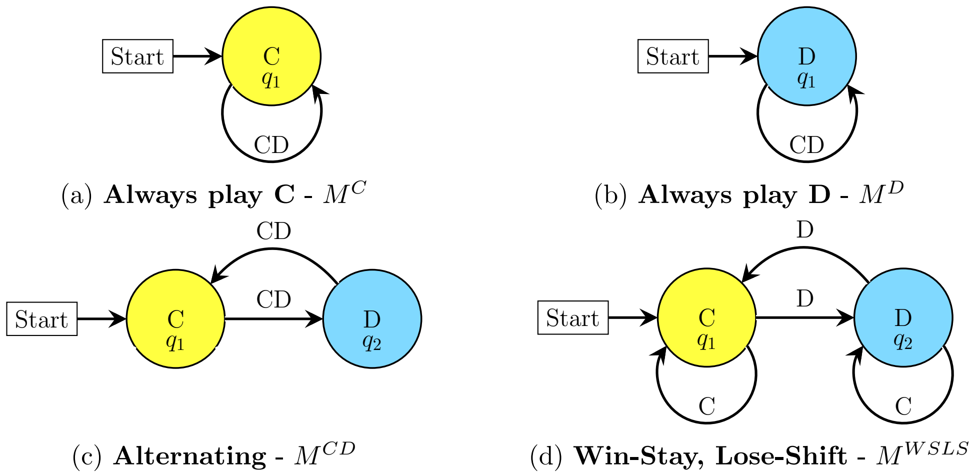


Figure 3.3: Important Two-State Automata.

3.4.2 Characterization of Equilibria

I give a characterization of the equilibria when players face stage-game payoffs presented in Figure 3.1. This game is a prisoner's dilemma when $L > 0$ with unique Nash equilibrium (D, D). When $L < 0$, the unique Nash equilibrium is (C, C), and it is no longer a prisoner's dilemma game.

I am interested in equilibria which are not heavily tied to the parameters of the game. I focus on robust equilibria. Say that $G(P_1, P_2)$ is the supergame where player i is subject to payoff matrix $P_i \in \mathbb{R}^{|S_1| \times |S_2|}$.

Definition 3.4.1 (Robust Equilibrium)**.** *Suppose two players play supergame $G(P_1, P_2)$ and have fixed signal functions r_i and error level ε. A pair of automata, (M_1, M_2), is a robust equilibrium of the supergame $G(P_1, P_2)$ if and only if there exists some $\mu > 0$ such that (M_1, M_2) is an equilibrium of all supergames $G(P_1', P_2')$ such that $\max_{s_i \in S_i, s_{-i} \in S_{-i}} |P_i'(s_i, s_{-i}) - P_i(s_i, s_{-i})| < \mu$.*

This equilibrium concept is a refinement of the Nash equilibrium concept defined in Definition 3.3.2. So every robust equilibrium is also a Nash equilibrium. The types of Nash equilibria that are not robust are only equilibria for a set of measure zero in the payoff space, and are therefore heavily tied to the parameters of the game. Robust equilibria are more universal than non-robust equilibria because they hold for a larger class of games. Therefore, they remain equilibria under small changes in the parameters. In the infinitely repeated PD game, there are at most two types of robust equilibria at any parameter pair.

Definition 3.4.2 (Payoff Equivalent Automata)**.** *Automata M_1 and M_2 are said to be payoff equivalent over set $\mathcal{M}$ if and only if,*

$$U_i\left(M_1, A, \varepsilon\right) = U_i\left(M_2, A, \varepsilon\right) \text{ for all } A \in \mathcal{M}, \text{ and all } \varepsilon \in (0, .5].$$

Automata M_1 and M_2 are said to be payoff equivalent only if they yield the same payoff against any other automata.

Theorem 3.4.3. *In the infinitely repeated PD game, when players have the simple signal function r_i^S and choose among the set of two-state automata, $\mathcal{M}^2$, there are only three types of robust equilibria:*

1. *$L < 0$ and M_i is payoff equivalent to M^C for $i = 1, 2$,*
2. *$L > 0$ and M_i is payoff equivalent to M^D for $i = 1, 2$, and*
3. *$-\left(1 - 2\varepsilon\right)^3 < L < \left(1 - 2\varepsilon\right)^3$ and $M_i = M^{WSLS}$ for $i = 1, 2$.*

The proof of this result is left to the appendix. Based on these regions, notice that there are at most two types of equilibria at any pair of payoff parameter and error level. The equilibrium regions are displayed in Figure 3.4. When both players play M^{WSLS}, high levels of cooperation occur in equilibrium. The dark shaded region in Figure 3.4 represents the area where both players playing M^{WSLS} is an equilibrium. Whenever the payoff parameter and error level is in this region, then both players playing M^{WSLS} is an equilibrium, and therefore high levels of cooperation are attainable in equilibrium.

When players play M^C or M^D, their strategies are unresponsive to the signals they receive. The only robust equilibrium where players' strategies are responsive to their signals is when both players play M^{WSLS}. What makes M^{WSLS} so special? When players are trying to cooperate, they must punish their opponents to deter deviations. When an incorrect signal is received, they may start to punish each other repeatedly. In order to sustain cooperation, they must somehow recoordinate their actions to start cooperating again after an incorrect signal has been received. Since recoordination

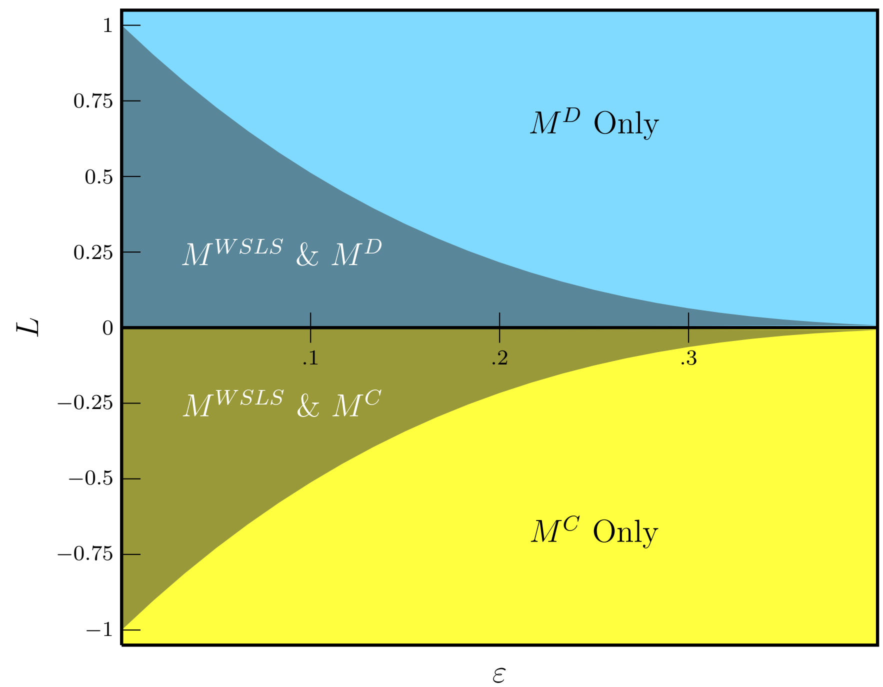


Figure 3.4: Equilibrium Regions from Theorem 3.4.3.

is typically inefficient, players want to recoordinate as quickly as possible after an incorrect signal is received. If both players play M^{WSLS}, this recoordination is efficient. Consider the following example which describes how this recoordination occurs. Suppose both players are playing M^{WSLS}, the sequence of states and signals is displayed in Table 3.1.

- **First Round** - Both players start in state q_1, both play C, both receive correct signals, and both transition back to state q_1 in round 2.

- **Second Round** - Both players again play C. Now, player 1 receives an incorrect signal that player 2 played D (error denoted with box around signal). Player 1 thinks player 2 played D and therefore moves to q_2. Player 1 received a correct signal, and therefore returns to state q_1.

- **Third Round** - Player 1 plays D and Player 2 plays C. Both receive correct signals. Player 1 sees that player 2 played C, so player 1 remains in q_2. Player 2 sees that player 1 played D, so player 2 switches to q_2.

- **Fourth Round** - Both players are in state q_2, both play D, both receive correct signals, and both switch states, and transition back to state q_1.

- **Fifth Round -** Both players are back in q_1, and continue to cooperate until another incorrect signal is received.

Round	1	2	3	4	5
Current State of M_1	q_1	q_1	q_2	q_2	q_1
Signal from M_2	C	D	C	D	C
Signal from M_1	C	C	D	D	C
Current State of M_2	q_1	q_1	q_1	q_2	q_1

Table 3.1: Recoordination of M^{WSLS}.

So in this example, the players start by cooperating in round one. The incorrect signal was received in round two. Player 2 plays D to confirm that an incorrect signal was received in round three. Both players play D in round four. In round five, both players are cooperating again. So after the incorrect signal is received, it only takes two rounds to recoordinate if no other incorrect signals are received. This efficient recoordination is one of the reasons why M^{WSLS} is an equilibrium strategy.

Another reason why M^{WSLS} is special is because it is not dominated by M^C or M^D for large regions of the parameter space. When players play $\left(M^{WSLS}, M^{WSLS}\right)$, the action pair (C, C) is played most of the time, so the players receive close to the cooperative payoff. In the system $\left(M^{WSLS}, M^C\right)$, the action pairs (C, C) and (D, C) are each played half the time. This is bad for player 2, because $u_2(D, C) = -L < u_2(C, C) = 1$ when $L > -1$. Playing M^C is only good for player 2 when L is sufficiently negative. In the system $\left(M^{WSLS}, M^D\right)$, action pairs (C, D) and (D, D) are each played half the time. Again this is not good for player 2 because $\frac{u_2(C,D)+u_2(D,D)}{2} = \frac{1+L}{2} \leq u_2(C, C)$ when $L < 1$. Playing M^D is only profitable for player 2 if L is sufficiently high. For medium ranges of L, M^{WSLS} is the best response to itself, because it receives the cooperative payoff most of the time.

This result does not depend on the prisoner's dilemma game. Similar results hold for a class of Battle of the Sexes games as well as a class of minimum-effort coordination games. In both of these cases, the only types of equilibria either are unresponsive to the signal of the other players action, or similar to M^{WSLS}. These results (Theorems B.2.18 and B.2.19) are left to the appendix.

3.4.3 General 2×2 Games

In the previous section, I showed that both players playing M^{WSLS} is an equilibrium for a large region of the parameter space when players play an infinitely repeated prisoner's dilemma game. In this section, I give conditions on stage-game payoffs, which ensure that $\left(M^{WSLS}, M^{WSLS}\right)$ is an equilibrium.

Theorem 3.4.4. *Suppose both players have simple signal functions r_i^S. If for $i = 1, 2$,*

1. $u_i(C, C) > u_i(C, D)$, *and*

2. $u_i(C, C) > \frac{u_i(D,C) + u_i(D,D)}{2}$;

then there exists some $\bar{\varepsilon} > 0$ such that $\left(M^{WSLS}, M^{WSLS}\right)$ is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$.

This result suggests that when errors are small $\left(M^{WSLS}, M^{WSLS}\right)$ is an equilibrium for a wide range of games. Figure 3.5 displays four 2×2 games that satisfy the desired properties.

- Figure 3.5(a) is a stag-hunt game with Pareto ranked pure strategy Nash equilibria (C, C) and (D, D). Both players playing M^{WSLS} leads to high levels of the Pareto superior equilibrium.

- Figure 3.5(b) is a chicken game with two pure strategy Nash equilibria (C, D) and (D, C), one preferred by each player. If both play M^{WSLS}, then the cooperative outcome (C, C) is possible, even though it is not one of the pure strategy Nash equilibria.

- Figure 3.5(c) is a Battle of the Sexes game with two pure strategy Nash equilibria (C, C) and (D, D). If both players play M^{WSLS} then the outcome (C, C) is frequently attained. Also consider, $M^{LSWS} = (\{C, D\}, \{2, 1\}, \{1, 2\})$. This strategy is the opposite of M^{WSLS} in that it stays in the same state when the other player plays D, and switches when the other player plays C. The theorem also confirms that both players playing M^{LSWS} is also an equilibrium in this case.

	C	D
C	4,4	1,3
D	3,1	3,3

(a) Stag-Hunt

	C	D
C	4,4	2,5
D	5,2	0,0

(b) Chicken

	C	D
C	4,3	0,0
D	0,0	3,4

(c) Battle of the Sexes

	C	D
C	4,3	2,4
D	1,2	3,1

(d) No Pure Equilibrium

Figure 3.5: 2×2 Games.

- Figure 3.5(d) is a game with no pure strategy equilibrium. However, both players playing M^{WSLS} leads to high levels of (C, C) in equilibrium.

So the simple strategy M^{WSLS} is an equilibrium for a variety of 2×2 games when errors are small.

3.4.4 Other Support

The strategy represented by automaton M^{WSLS} has been studied before. The majority of work done on this strategy focuses on biological applications. Nowak and Sigmund (1993) run evolutionary simulations on probabilistic memory one strategies. Memory one strategies are those which only respond to the previous period of play, similar to the two-state case. Their simulations are more general than my two-state result because they allow for probabilistic transitions. Nevertheless, the prevailing strategy in their simulation is the deterministic M^{WSLS} strategy.

More recently, Imhof, Fudenberg, and Nowak (2007) use stochastic evolutionary game dynamics to study the evolution of four strategies, $M^C, M^D, M^{TFT}, M^{WSLS}$. When only M^C, M^D, and M^{WSLS} are considered, they find some payoff threshold which determines which strategy is selected. Below this threshold M^D is selected while above this threshold M^{WSLS} is selected. When M^{TFT} is added to the three other strategies, they again find a threshold, but this time it is lower, meaning that M^{TFT} strengthens M^{WSLS}.

The prediction from my two-state model is that the only equilibria in the infinitely repeated prisoners dilemma game ($L > 0$) are M^D or M^{WSLS}. Experiments with human subjects playing repeated prisoners dilemma games have often tried to identify subjects playing tit-for-tat (Dal Bó and Fréchette 2008). Tit-for-tat typically fits the data well. One of the reasons why tit-for-tat fits the data well is that human subjects tend to always play C or always play D, both of which are supported by tit-for-tat. The predictions of my model also support this behavior. However, there is one key difference between M^{TFT} and M^{WSLS} or M^D that allows us to identify which strategies the subjects are playing.

To identify a strategy, look at the play of both players in round t, and then see the responses in $t+1$. If player 1 is playing M^{TFT} and both players play C in round t, then player 1 should play C in round $t+1$. M^{WSLS} provides the same prediction, that both players playing C leads to player 1 playing C. M^{TFT} and M^{WSLS} again share a prediction if player 1 plays C and player 2 plays D in round t. Both predict that D will be played in round $t+1$. If both players play D in round t, then M^{TFT} and M^D both predict that D is played in the following round. So far the predictions of M^{TFT} have matched the prediction of M^{WSLS} or M^D. The final combination is where they differ. If player 1 plays D and player 2 plays C in round t, then M^{TFT} predicts that player 1 will play C in the next round. Conversely, both M^D and M^{WSLS} predict that player 1 continues to play D in the next round. This provides a testable prediction: if player 1 plays D and player 2 plays C, then player 1 will play C in the next round if he is using M^{TFT}, and will play D in the next round if he is using M^{WSLS} or M^D.

Wedekind and Milinski (1996) run experiments that examine whether players play M^{TFT} or M^{WSLS}. They find that 70% of players can be classified a playing M^{WSLS} in a variety of treatments of repeated prisoner's dilemma game. Their experiments use psuedoplayers which use predetermined strategies. This allows them to focus on the situation of interest. To classify the strategies of players, they focus on the situation where player 1 plays D and player 2 plays C in round t. If player 1 plays C more in round $t+1$, then he is classified as playing M^{TFT}. If player 1 plays D more in round $t+1$, then he is classified as playing M^{WSLS}. These experimental results suggest that players are playing M^{WSLS}.

3.5 Unrestricted Automata

In this section, I examine the case where players can select any automata with any finite number of states to represent their strategies. I first introduce absorbing classes and communicating classes which are necessary to understand the equilibrium characterization. Then, I reanalyze the motivating example which shows the importance of absorbing and communicating classes. Finally, I give the necessary and sufficient conditions on equilibrium structure for small error levels.

3.5.1 Absorbing Classes

An absorbing class of an automaton M is a region that the automaton can get stuck in when the opponent plays a sequence of actions repeatedly. If player 1 plays automaton M_1, then player 2 wants to play a strategy which ensures that M_1 gets stuck in a high payoff absorbing class rather than a low payoff absorbing class. Formally,

Definition 3.5.1 (Absorbing Class). *Given automaton $M = (Q, q^0, f, \tau)$, an absorbing class, denoted by $a(M) = \{\boldsymbol{q}, \boldsymbol{s}\}$, where $\boldsymbol{q} = q_1, \ldots, q_n$ is a sequence of states, and $\boldsymbol{s} = s_1, \ldots, s_n$ is a sequence* *of signals, such that*

$$\tau\left(q_k, s_k\right) = \begin{cases} q_{k+1} & k < n \\ q_1 & k = n. \end{cases}$$

So when automaton M is in state q_1, and the opponent plays sequence $\boldsymbol{s}$ repeatedly, then M will loop through the sequence of states $\boldsymbol{q}$ repeatedly. The length of an absorbing class is the length of the sequences of actions and states, $|a| = n$. When automaton M is looping through states $\boldsymbol{q}$ and the opponent is playing actions $\boldsymbol{s}$, then the players are playing a fixed sequence of joint actions repeatedly. The payoff for an absorbing class is the average payoff per round while in this absorbing class,

$$U_i^{AC}\left(a\right) = \frac{1}{|a|}\sum_{k=1}^{|a|} u_i\left(f\left(q_k\right), s_k\right).$$

The set of all possible absorbing classes for automaton $M = \left(Q, q^0, f, \tau\right)$ is infinite. However there exists a payoff-maximal absorbing class for player i, denoted by $a_i^*(M)$, with $|a_i^*(M)| \leq |Q|$. This result, presented in Lemmas B.2.2 and B.2.3, is left to the appendix. The idea for the proof is that if a payoff-optimal absorbing class travels through the same state twice, then there must be a smaller payoff-optimal absorbing class. Therefore, given any payoff-optimal absorbing class, the length can be reduced by eliminating states that appear more than once, until it has length less than or equal to $|Q|$. This finite length optimal absorbing class is used to construct a best response automaton. In equilibrium, each player must spend almost all of the repeated game in the optimal absorbing class of their opponents automaton.

3.5.2 Communicating Class

Once both players have selected automata $M_1 = \left(Q_1, q_1^0, f_1, \tau_1\right)$ and $M_2 = \left(Q_2, q_2^0, f_2, \tau_2\right)$, the pair of automata (M_1, M_2) forms a *system* which can be represented with a finite Markov chain $X\left(M_1, M_2, \varepsilon\right)$. Each state of the Markov chain corresponds to a pair of automaton-states, one from each automaton. For example, the situation where M_1 is in state q_1 and M_2 is in state q_2 is represented by one state of the Markov chain. The starting state of the Markov chain represents the situation where both automata are in their initial states, M_1 in q_1^0 and M_2 in q_2^0. Based on the signal functions r_i, the Markov chain has $n \leq |Q_1|\,|Q_2|$ states, one corresponding to each pair of automaton states that are reachable from the initial states with positive probability for any $\varepsilon > 0$. These states are denoted by $x_1, \ldots, x_n$.

Let $q_i(x)$ be the current state of automaton M_i when the Markov chain is in state x. Automaton M_i moves from state $q_i(x_a)$ to $q_i(x_b)$ with probability,

$$\mathbb{P}\left(M_i, q_i\left(x_a\right), q_i\left(x_b\right), \varepsilon\right) = \sum_{s_i | \tau\left(q_i\left(x_a\right), s_i\right) = q_i\left(x_b\right)} r_i\left(s_i, f_{-i}\left(q_{-i}\left(x_a\right)\right), \varepsilon\right).$$

In words, the term inside the sum is the probability that player i receives a signal that the other player played action s_i when the other player actually played action $f_{-i}\left(q_{-i}(x_a)\right)$. This term is then summed over all actions s_i which take automaton M_i from state $q_i(x_a)$ to $q_i(x_b)$. The Markov chain is therefore defined by the probability that M_1 moves from $q_1(x_a)$ to $q_1(x_b)$ and M_2 moves from $q_2(x_a)$ to $q_2(x_b)$,

$$X\left(M_1, M_2, \varepsilon\right)\left(x_a, x_b\right) = \mathbb{P}\left(M_1, q_1\left(x_a\right), q_1\left(x_b\right), \varepsilon\right) \mathbb{P}\left(M_2, q_2\left(x_a\right), q_2\left(x_b\right), \varepsilon\right). \tag{3.2}$$

The starting point of this Markov chain is state x^0 such that $q_1\left(x^0\right) = q_1^0$ and $q_2\left(x^0\right) = q_2^0$. When the signals are perfect, the Markov chain $X\left(M_1, M_2, 0\right)$ is deterministic. Each state leads to another state with probability 1. When the signals are imperfect, the Markov chain $X\left(M_1, M_2, \varepsilon\right)$ is not necessarily deterministic and any state may lead to multiple different states with varying probabilities.The realizations of the Markov chain are denoted by $x^1, x^2, \ldots$.

Definition 3.5.2 (Communicating Class)**.** *A communicating class of the system $X\left(M_1, M_2, \varepsilon\right)$ is a set of states $A \subseteq X\left(M_1, M_2, \varepsilon\right)$ that satisfies,*

- $\left(X\left(M_1, M_2, 0\right)(x, y)\right)^n = 0$ *for all* $x \in A, y \notin A, n > 0$.
- $\left(X\left(M_1, M_2, 0\right)(x, y)\right)^n > 0$ *for all* $x, y \in A$ *and for some* $n > 0$.

A communicating class is a set of absorbing states. Once the Markov chain enters a communicating class, it can only leave if a player receives an incorrect signal. When no erroneous signals are received, the Markov chain deterministically loops through the states in the communicating class. The payoff of a communicating class is defined to be the average payoff over this loop,

$$U_i^{CC}\left(A_k\right) = \frac{1}{|A_k|} \sum_{x \in A_k} u_i(x), \tag{3.3}$$

where $u_i\left(x\right) = u_i\left(f_1\left(q_1\left(x\right)\right), f_2\left(q_2\left(x\right)\right)\right)$ is the payoff for player i in state x. This definition gives the average payoff in the communicating class when signals are correct, and is used when giving necessary and sufficient conditions in the finite-state case.

3.5.3 Calculating Payoffs

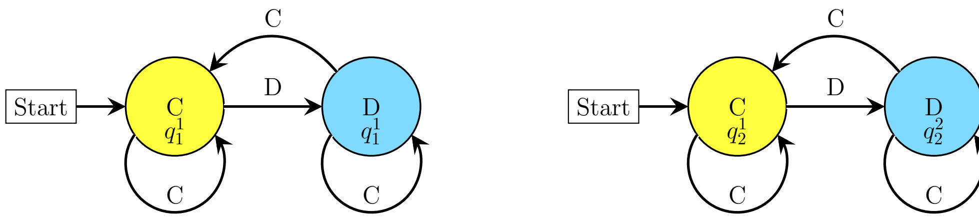
Representing the system as a Markov chain allows me to calculate the payoffs for a given pair of automata using only the stationary distribution of the Markov chain. By Lemma B.2.1, the Markov chain $X\left(M_1, M_2, \varepsilon\right)$ is irreducible for all $\varepsilon > 0$, and hence has a unique stationary distribution, $\pi\left(M_1, M_2, \varepsilon\right)$.

Lemma 3.5.3. *Suppose players play automata M_1 and M_2. The average payoff for the infinitely repeated game is equal to*

$$U_i\left(M_1, M_2, \varepsilon\right) = \sum_{x_k \in X\left(M_1, M_2, \varepsilon\right)} \pi\left(M_1, M_2, \varepsilon\right)\left(x_k\right) u_i\left(x_k\right),$$

where $\pi\left(M_1, M_2, \varepsilon\right)\left(x_k\right)$ is the term of the stationary distribution corresponding to state x_k, and $u_i\left(x_k\right)$ is the payoff for player i state x_k.

Lemma 3.5.3 implies that only the stationary distribution of the system and vector of utilities for the corresponding states are needed to find the limit of means for a pair of automata. The idea behind the proof is that the frequency of time the Markov chain spends in a state converges to the stationary distribution by the law of large numbers. The proof of this lemma is left to the appendix.



(a) Player 1 plays M^{TFT} (b) Player 2 plays M_2^{TFT}

Figure 3.6: Automata for absorbing class example.

3.5.4 Tit-For-Tat Absorbing Class Example

To better understand the absorbing classes, communicating classes, and where the analysis is headed, it helps to give an example. Assume the players are playing the PD-game presented in Figure 3.1. As the motivating example shows, tit-for-tat can get into trouble by getting caught in a suboptimal region when signals are imperfect. Here, I elaborate on the motivating example using the notation introduced above. This example helps to show why these suboptimal regions are a problem, and how to make automata more robust to these imperfect signals.

Suppose player 1 plays automaton M^{TFT} displayed in Figure 3.6(a). This is a two-state automaton with states q_1 and q_2 which represents the tit-for-tat strategy. The optimal absorbing class of M^{TFT} for player 2 is,

$$a_2^*\left(M^{TFT}\right) = \{\{q_1\}, \{C\}\}.$$

In this absorbing class, M^{TFT} is in state q_1, player 2 plays C, and M_1 returns to q_1. Therefore, the average payoff in this absorbing class for player 2 is $U_2^{AC}(a_2^*(M^{TFT})) = 1$.

Consider two other absorbing classes,

$$a^{CD}(M^{TFT}) = \{\{q_1, q_2\}, \{D, C\}\} \text{ and } a^D(M^{TFT}) = \{\{q_2\}, \{D\}\}.$$

The respective payoffs are $U_2^{AC}(a^{CD}) = 1/2$ and $U_2^{AC}\left(a^D\right) = 0$.

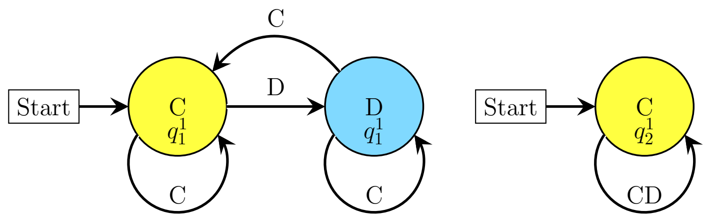
Ideally, player 2 would like to play an automaton which spends most of the supergame in absorbing class $a_2^*(M^{TFT})$, and does not get stuck in $a^{CD}(M^{TFT})$ or $a^D(M^{TFT})$. When an incorrect signal is received, the best response automaton should be able to find its way back to the payoff-optimal absorbing class without getting stuck in a suboptimal absorbing class.

Now, suppose that player 2 also plays M^{TFT} with states q_1 and q_2 as displayed in Figure 3.6(b). Suppose each player has the simple signal function r_i^S from (3.1). The Markov chain for the system, $X\left(M^{TFT}, M^{TFT}, \varepsilon\right)$, has states $\left\{x^{CC}, x^{CD}, x^{DC}, x^{DD}\right\} = \{q_1q_1, q_1q_2, q_2q_1, q_2q_2\}$,

$$X\left(M^{TFT}, M^{TFT}, \varepsilon\right) = \begin{array}{c} x^{CC} \\ x^{CD} \\ x^{DC} \\ x^{DD} \end{array} \begin{bmatrix} (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ \varepsilon(1-\varepsilon) & \varepsilon^2 & (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) \\ \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 & \varepsilon^2 & \varepsilon(1-\varepsilon) \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 \end{bmatrix}.$$

When ε is small, and the system is in state x^{CC}, then it returns to x^{CC} with probability $(1-\varepsilon)^2$ (only leaves with an incorrect signal). When the system is in state x^{CD}, then it goes to state x^{DC} and vice-versa, unless an incorrect signal is received. When in x^{DD}, the system remains in x^{DD}, unless an incorrect signal is received. So this system has three communicating classes: $A^C = \left\{x^{CC}\right\}, A^{CD} = \left\{x^{CD}, x^{DC}\right\}$, and $A^D = \left\{x^{DD}\right\}$. Once the system has entered a communicating class, it remains in that class until an incorrect signal is received. Note that $U_2^{CC}\left(A^C\right) = U_2^{AC}\left(a_2^*\left(M_1\right)\right)$, so when the system is in communicating class A^C, automaton M_1 is in player 2's optimal absorbing class. So player 2 wants the system to spend as much time in communicating class A^C as possible.

Since automaton M^{TFT} starts in state q_1, the system starts in state x^{CC}. When $\varepsilon = 0$, the starting point of the system matters, and the stationary distribution is $\pi\left(M^{TFT}, M^{TFT}, 0\right) = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array}\right]$. This means the system always stays in x^{CC}, and both players receive payoffs equal to their optimal absorbing class. Since neither player can do better, this is an equilibrium when $\varepsilon = 0$. For any $\varepsilon > 0$, the starting point no longer matters, and the unique stationary distribution is $\pi\left(M^{TFT}, M^{TFT}, \varepsilon\right) = \left[\begin{array}{cccc} 1/4 & 1/4 & 1/4 & 1/4 \end{array}\right]$. This means for any positive error level, the system



(a) Player 1 plays M^{TFT} (b) Player 2 plays M_2^C

Figure 3.7: Automata for absorbing class example.

spends one-quarter of the time in A^C and A^D, and one-half the time in A^{CD}, which yields payoff

$$U_2\left(M^{TFT}, M^{TFT}, \varepsilon\right) = \frac{1}{4}U_2^{CC}\left(A^C\right) + \frac{1}{2}U_2^{CC}\left(A^{CD}\right) + \frac{1}{4}U_2^{CC}\left(A^D\right) = \frac{1}{2} < U_2^{AC}\left(a_2^*\left(M^{TFT}\right)\right).$$

Since the system is in suboptimal absorbing classes A^{CD} or A^D three-quarters of the time, the payoff is less than the payoff of the optimal absorbing class. Player 2 spends significant time in a suboptimal absorbing class, even in the limit as the errors go to zero. Player 2 gets higher payoff playing an automaton that does not get caught in these suboptimal absorbing classes.

There is a discontinuity in the stationary distribution at $\varepsilon = 0$.

$$\pi\left(M^{TFT}, M^{TFT}, 0\right) = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array}\right] \neq \frac{1}{4}\left[\begin{array}{cccc} 1 & 1 & 1 & 1 \end{array}\right] = \lim_{\varepsilon \to 0} \pi\left(M^{TFT}, M^{TFT}, \varepsilon\right).$$

Since the payoffs for a pair of automata depend on the stationary distribution, a discontinuity like this in the stationary distribution, also leads to a discontinuity in the payoffs,

$$U_i\left(M^{TFT}, M^{TFT}, 0\right) = 1 \neq \frac{1}{2} = \lim_{\varepsilon \to 0} U_i\left(M^{TFT}, M^{TFT}, \varepsilon\right).$$

Because of this discontinuity, some of the equilibria under perfect monitoring may fail to be equilibria when monitoring becomes imperfect. If the stationary distribution does not have this discontinuity, then neither do payoffs.

Next, suppose that player 1 still plays the tit-for-tat automaton M^{TFT}. Player 2 now plays the "always play C" automaton M^C, with state q', as displayed in Figure 3.7(b). Since M^C only has one state, the system only has two states $\{x^{CC}, x^{DC}\} = \{q_1 q', q_2 q'\}$. The corresponding Markov chain is,

$$X(M_1, M_2, \varepsilon) = \begin{matrix} x^{CC} \\ x^{DC} \end{matrix} \begin{bmatrix} 1-\varepsilon & \varepsilon \\ 1-\varepsilon & \varepsilon \end{bmatrix}.$$

When this system is in x^{CC}, it returns to x^{CC} unless an incorrect signal is received. State x^{DC} also leads to x^{CC} unless an incorrect signal is received. So now there is only one communicating class in this system, $A^C = \{x^{CC}\}$. There is also a transient class, $T^C = \{x^{DC}\}$, which leads to communicating class A^C. As before, the payoff of this communicating class is equal to the payoff for player 2's optimal absorbing class, $U_2^{CC}(A^C) = U_2^{AC}(a_2^*(M_1))$. However, now there are no suboptimal absorbing classes.

When the signals are perfect, the stationary distribution of the system is $\pi(M_1, M_2, 0) = \begin{bmatrix} 1 & 0 \end{bmatrix}$. This means that the system spends all of the time in x^{CC}, and yields payoff $U_2(M^{TFT}, M^C, 0) = 1$. It turns out that M^C is a best response to M^{TFT} when the signals are perfect. When $\varepsilon > 0$, the stationary distribution becomes $\pi(M_1, M_2, \varepsilon) = \begin{bmatrix} 1-\varepsilon & \varepsilon \end{bmatrix}$, which yields payoff

$$U_2(M^{TFT}, M^C, \varepsilon) = (1-\varepsilon) U_2(A^C) + \varepsilon(-L) = 1 - \varepsilon(1+L).$$

Now player 2 spends $(1-\varepsilon)$ in the optimal absorbing class and only ε in the transient class. Player 2 never gets caught in some suboptimal absorbing class. As ε approaches zero, player 2 is almost always in the optimal absorbing class, and cannot do better.

Unlike the previous case, there is no longer a discontinuity in the stationary distributions [4],

$$\pi(M^{TFT}, M^C, 0) = \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} = \lim_{\varepsilon \to 0} \pi(M^{TFT}, M^C, \varepsilon).$$

[4] In the literature on perturbed Markov chains, there are two types of perturbations: *regular perturbations* and *singular perturbations*. In a regular perturbation, the stationary distribution is continuous as the error increases from zero. A singular perturbation is defined as a perturbation which changes the ergodic structure of the matrix, meaning that multiple communicating classes may be combined. Singular perturbations typically do not have continuous stationary distributions when moving the errors away from zero. $X(M^{TFT}, M^{TFT}, \varepsilon)$ is a singular perturbation while $X(M^{TFT}, M^C, \varepsilon)$ is a regular perturbation (different than regular perturbation defined in this paper).

Again, since the payoffs are just a linear function of the stationary distribution, continuous stationary distributions lead to continuous payoffs,

$$U_2\left(M^{TFT}, M^C, 0\right) = 1 = 1 = \lim_{\varepsilon\to 0} U_2\left(M^{TFT}, M^C, \varepsilon\right).$$

Since M^C was a best response with perfect signals, and the payoffs are continuous, it remains a best response in the limit as the probability of getting an incorrect signal goes to zero.

To summarize, if player 1 plays M_1, then player 2 wants to play an automaton such that automaton M_1 spends most of the time in player 2's optimal absorbing class $a_2^*(M_1)$ and does not get caught in suboptimal absorbing classes. Suboptimal absorbing classes can be detrimental to payoffs, even in the limit as the probability of an incorrect signal goes to zero.

3.5.5 Necessary and Sufficient Conditions

In this section, I provide the necessary and sufficient conditions for equilibria in the finite-state case. As the above example shows, in equilibrium, players cannot spend significant amounts of time in communicating classes that do not yield the optimal absorbing class payoff. If all communicating classes yield the same payoff as the optimal absorbing class for each player, then it is an equilibrium. However, it is possible to have communicating classes which do not yield the optimal absorbing class payoff in equilibrium as long as the time spent in these communicating classes is significantly less than time spent in other communicating classes.

To formalize these conditions, we must understand how some communicating classes are more robust to incorrect signals than others. The system may exit some communicating classes with only one incorrect signal, while others may require many more incorrect signals. The system visits those communicating classes that are most robust to incorrect signals almost all the time as the probability of error goes to zero.

Definition 3.5.4 (Prevalent Communicating Class)**.** *A communicating class A of the matrix $X\left(M_1, M_2, \varepsilon\right)$*

is a prevalent communicating class if

$$\lim_{\varepsilon \to 0} \pi \left(M_1, M_2, \varepsilon\right)(x) > 0 \text{ for some } x \in A.$$

A prevalent communicating class is a set of states that the Markov Chain $X(M_1, M_2, \varepsilon)$ visits with positive probability in the limit as the error goes to zero. When ε is small, the system spends almost all the time in the prevalent communicating classes.

Next, these results hold for a wider class of signal functions, as defined below,

Definition 3.5.5 (Regular Signal Function)**.** *A signal function $r_i : S_{-i} \times S_{-i} \times [0, .5] \to [0, 1]$ is said to be regular if the following conditions hold.*

1. $\lim_{\varepsilon \to 0} r_i(s_i, s_j, \varepsilon) = \begin{cases} 1 & s_i = s_j \\ 0 & s_i \neq s_j \end{cases}$,
2. $r(s_i, s_j, \varepsilon) > 0$ *for all* $\varepsilon \in (0, .5]$ *and all* $s_i, s_j \in S_{-i}$,
3. $\exists n \geq 0$ *such that* $0 < \lim_{\varepsilon \to 0} \varepsilon^{-n} r(s_i, s_j, \varepsilon) < \infty$ *for all* $s_i, s_j \in S_{-i}$.

It is clear that the simple signal function, r_i^S from (3.1), is a regular signal function. There are also more complex signal functions that satisfy this as well.

For the finite-state results, I restrict the set of finite automata to those which are finite, strongly connected, and reduced. This set is denoted by $\mathcal{M}_i^R$. All equilibria over this set are also equilibria over the set of all finite automata. For more details see Appendix B.1. With this notation, I introduce the main results for the finite-state case.

Theorem 3.5.6 (Necessity)**.** *Suppose players play supergame G with regular signal function r_i, and play automata $M_i \in \mathcal{M}_i^R$ represented by Markov chain $X(M_1, M_2, \varepsilon)$. If there exists some $\bar{\varepsilon} > 0$ such that (M_1, M_2) is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$, then for all prevalent communicating classes A_k, $U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i}))$.*

These conditions say that, for small error levels, each prevalent communicating class must yield the optimal absorbing class payoff for each player. Since almost all time is spent in the prevalent communicating classes when the errors are small, the system must spend almost all the time in good regions and not get caught in bad regions in equilibrium.

To prove the necessary conditions, I show that if the necessary conditions are not satisfied, then it is always possible to construct an automaton M_2' such that for some $\bar{\varepsilon} > 0$,

$$U_2(M_1, M_2, \varepsilon) < U_2(M_1, M_2', \varepsilon) \text{ for all } \varepsilon \in (0, \bar{\varepsilon}).$$

So M_2' is a better response than M_2 to automaton M_1. This means that (M_1, M_2) is not an equilibrium if the desired properties are not satisfied. I show that such an automaton M_2' exists in the following lemma.

Lemma 3.5.7. *Given automaton $M_1 \in \mathcal{M}^R$ with n states, and any absorbing class $a(M_1)$, there exists automaton M_2 such that for all communicating classes, A_k, of the system $X(M_1, M_2, \varepsilon)$,*

$$U_2^{CC}(A_k) = U_2^{AC}(a(M_1)).$$

The proof of the Lemma is left to the appendix. To prove this I construct automaton M_2 with the desired properties. Automaton M_2 contains three regions. The first region is called the *absorbing region.* As long as no incorrect signals are received, M_2 remains in this region when M_1 is in the desired absorbing class $a(M_1)$. When an incorrect signal is received by either player, there is a chance that automaton M_1 will leave the states of $a(M_1)$. When this happens, player 2 becomes confused about the current state of M_1, and must try to make inferences about current state. Player 2 wants to get back to the states of $a(M_1)$ without getting caught in another suboptimal absorbing class. To do this, player 2 plays what is called a *homing sequence.* This homing sequence is a fixed sequence of actions, which based on the output, determines the current state of M_1. Once the current state of M_1 has been identified, the automaton enters the *resynchronization region.* When entering this region, M_2 knows the current state of M_1. M_2 then plays a sequence of actions which resynchronizes the automata after which automaton M_1 returns to the states of the desired absorbing class, and automaton M_2 returns to the absorbing region. Automaton M_1 remains in the states of the desired absorbing class until an incorrect signal is received. To better understand the construction, I provide an example in Appendix B.3.2.

Next, I give the sufficient conditions for the structure of equilibrium automata. Let $\mathcal{M}^{SPM}(M_i)$ be the set of all automata $M_{-i} \in \mathcal{M}_{-i}^R$ such that all prevalent communicating classes of $X(M_i, M_{-i}, \varepsilon)$, A_k, yield the optimal absorbing class payoff, $U_i^{CC}(A_k) = U_i^{AC}(a_i^*(M_{-i}))$ for $i = 1, 2$. This is the set of all automata that when paired with M_i yield the optimal absorbing class payoff in all prevalent communicating classes.

Theorem 3.5.8 (Sufficiency)**.** *Suppose players play supergame G with regular signal function r_i, and play automata $M_i \in \mathcal{M}_i^R$ represented by Markov chain $X(M_1, M_2, \varepsilon)$. If*

1. *for all prevalent communicating classes A_k, $U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i}))$, and*

2. $\frac{\partial U_i(M_1, M_2, 0)}{\partial \varepsilon} = \sup_{M \in \mathcal{M}^{SPM}(M_{-i})} \frac{\partial U_i(M_i, M, 0)}{\partial \varepsilon}$*;*

then there exists some $\bar{\varepsilon} > 0$ such that (M_1, M_2) is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$.

This theorem provides sufficient conditions for equilibrium automata in the finite-state case when errors are sufficiently small. The first condition requires that all prevalent communicating classes yield the optimal absorbing class for both players. Since the system spends almost all the time in prevalent communicating classes and almost no time in the other states, this formalizes the intuition that the system cannot get stuck in suboptimal regions for long periods of time. The second condition requires that out of all $M \in \mathcal{M}^{SPM}(M_i)$, the player must select the one that yields the highest marginal utility at zero. The two conditions together are then necessary and sufficient for equilibrium for sufficiently small errors. The proof of the sufficient conditions is left to the appendix.

3.6 Discussion

3.6.1 Literature Review

There has been a lot of work done examining imperfect monitoring in repeated games. The different models of imperfect monitoring all share the common theme that the players must recoordinate after an error is made. When there is some common knowledge among players, recoordination is relatively easy. Models involving imperfect public monitoring (Fudenberg, Levine, and Maskin 1994) as well as models of imperfect private monitoring with communication (Compte 1998, Kandori and Matsushima 1998, Obara 2009) are able to obtain the folk theorem. This common knowledge allows for relatively easy recoordination.

When there is no common knowledge, as in the imperfect private monitoring case, coordination becomes more difficult because players are not able to condition their strategies on a common knowledge signal and therefore must make inferences about the actions of their opponents. There are two main approaches to the study of equilibria under imperfect private monitoring: the belief-based approach and the belief-free approach. In the belief based approach players must make statistical inferences about the history of actions. The inferences quickly become difficult, and therefore results here typically require signals to be highly correlated (i.e. almost public) in order to obtain meaningful results (Bhaskar and Obara 2002, Mailath and Morris 2002). In the belief-free approach, strategies are constructed to ensure that beliefs are irrelevant. The prevailing equilibrium strategies are complex, and require that players are indifferent over a set of actions (Piccione 2002, Ely, Horner, and Olszewski 2005, Yamamoto 2009). If the payoffs are perturbed slightly, then these strategies fail to remain equilibria (Bhaskar 2000). So previous models of imperfect private monitoring are heavily dependent on either payoffs or signal structure. In this paper, the players use strategies represented by finite automata to represent their strategies which allows for a simpler representation of the beliefs.

There has also been work examining repeated games when players have bounds on memory. Lehrer (1988) and Sabourian (1998) look at models where players have bounded recall and perfect monitoring, while Cole and Kocherlakota (2005) examine a model of bounded recall with imperfect public monitoring. These results typically examine the effect of memory length on possible outcomes. Others have examined models where players select finite automata as their strategies. Using finite automata to represent strategies was first suggested by Aumann (1981). Since then, applications have included looking at finitely repeated games (Neyman 1985), assuming players have some exogenous cost of complexity (more states more costly) on their strategies (Rubinstein 1986, Abreu and Rubinstein 1988), or examining the evolutionary stability of such strategies (Miller 1996, Ioannou 2009). It is important to note that not every finite automaton strategy can be represented with a bounded memory strategy, but every bounded memory strategy can be represented as an automaton (Cole and Kocherlakota 2005).

3.7 Conclusion

The paper started with the question: Does a model of repeated interactions with boundedly rational agents lead to a smaller set of outcomes in equilibrium? When player's are limited to two-state automata, the number of outcomes in equilibrium is small (Theorem 3.4.3). Importantly, in an infinitely repeated prisoner's dilemma game, high levels of cooperation are still possible in equilibrium, even when agents cannot perfectly monitor their opponents and have no common knowledge public signal with which to recoordinate. The important strategy used is called "Win-Stay, Lose-Shift". If both players play this strategy, when cooperation breaks down, the players are able to quickly recoordinate and get back to cooperation without getting stuck in conflict for long periods. I show that WSLS holds for a variety of 2×2 games as well (Theorem 3.4.4). So when restricted to two-state automata, the number of equilibrium outcomes is small, and the predictions match the behavior of human subjects in the laboratory.

When I remove the restriction of two-state automata, the analysis becomes more difficult. In this case, I am able to provide necessary and sufficient conditions on equilibrium structure for small error levels (Theorems 3.5.6 and 3.5.8). The results show that in equilibrium players must play strategies which are able to cooperate without getting stuck in long periods of conflict. However, the implications of these conditions on the set of equilibrium outcomes remains an open question.

There are many extensions for this work. First, a better understanding of the effect of the necessary and sufficient conditions on outcomes. It is possible that for even small errors and finite-state strategies, the set of outcomes could still be small compared to the folk theorem. Also there is more work to be done examining what happens for larger errors when players can use finite-state automata as their strategies. In addition, more experiments with human subjects to further verify that these strategies are actually played in the lab.

There are also some more broad extensions. Assuming that players use finite automata as their strategies is assuming that they are classifying the infinite set of repeated game histories into a finite set of groups. It would be interesting to examine more general classification systems that would allow players to have more general groupings of their histories, rather than just those that can be represented with a finite automaton. Also, this paper only focuses on the equilibria, but there may be some learning that takes place to get to these equilibria. If we assume that players use automata to represent their strategies, there are a number of interesting learning dynamics that the players could use to learn to play certain strategies.

Chapter 4

Hysteresis in Coordination Games

4.1 Introduction

The goal of game theory is to predict how individuals behave when faced with strategic interactions. When these strategic interactions have multiple equilibria, it may become difficult to predict which equilibrium the players are going to play. There has been much work focusing on refining the set of equilibria to allow for sharper predictions when the game has multiple equilibria. This paper examines the effect of hysteresis on equilibrium selection. A system exhibits hysteresis if the outcome of a game depends on the history leading up to that game. I show that under certain conditions this hysteresis is possible in coordination games. In addition, laboratory experiments confirm the hysteresis hypothesis.

The fundamental tool used to make predictions about behavior in these games is the Nash equilibrium. When there is only one Nash equilibrium, the theory provides a solid prediction about how the individuals will play a certain game. However, in many games, there are multiple Nash equilibria, which make it difficult to make predictions because it is not apparent which equilibrium the players will select. Therefore, a lot of work has been done on better understanding which equilibrium will be selected in these games with multiple equilibria.

Harsanyi and Selten (1988) suggest two different methods for selecting equilibria in games with multiple equilibria: "payoff dominance" and "risk dominance". In certain situations, these two selection criteria may conflict, meaning the payoff dominant equilibrium is not the risk dominant equilibrium. When there is conflict, Harsanyi and Selten (1988) suggest that payoff dominance should be used instead of risk dominance. Since this however, a growing amount of support for the risk dominant equilibrium has emerged. Kandori, Mailath, and Rob (1993) and Young (1993) examine evolutionary stability of strategies that are subject to small mutations. They find that the equilibria that survive this evolutionary process are the risk dominant equilibria. Carlsson and Damme (1993) look at a model of incomplete information where the game is randomly chosen from a class of games, and find that the risk dominant equilibrium is selected. This paper shows that there is hysteresis, which means that different equilibria can be selected in the same game if there is different history leading to each of these games.

Hysteresis has been shown to hold in a wide variety of physical settings including magnetism and elasticity, but has also been observed in economics. Blanchard and Summers (1986) present a model in which the natural unemployment rate exhibits hysteresis in the presence of shocks. Employers make employment decisions in advance with the goal of maintaining steady employment in expectation. Employment shocks change these expectations and lead to more permanent changes in the natural unemployment rate. Baldwin (1988) shows that overvaluation of the dollar lead to hysteresis in United States import prices. Dixit (1989) examines entry of Japanese firms into the US market based on exchange rate fluctuations and finds that due to sunk costs, firm may remain in US even after the favorable exchange rate fluctuation has subsided. Nyberg (1997) examines an evolutionary model of honesty, and finds that once a society loses its honesty, hysteresis makes it difficult to reestablish. This paper focuses on coordination games, and finds that hysteresis occurs in these settings as well.

Coordination games have been a common area of study in the game theory literature. These are games in which individuals benefit from coordinating together, but are unable to communicate or enforce contracts, which can make coordination difficult, even if it is in both players best interest. By definition, all coordination games have multiple equilibria, so the question of equilibrium selection has been the main focus of the research for these coordination games.

Cooper, DeJong, Forsythe, and Ross (1990) show that Pareto dominant Nash equilibria are not always chosen in two player coordination games. They also find that payoffs of strictly dominated strategies play a large role in equilibrium selection. Van Huyck, Battalio, and Beil (1990) perform similar experiments as Cooper, DeJong, Forsythe, and Ross (1990), but consider up to 16 subjects per round, rather than just two. They find a large effect of group size on the final outcome of the coordination game, and also that Pareto dominant outcomes are very difficult to obtain when there is risk involved with large groups (16 subjects). The above papers consider equilibrium selection in coordination games. However, they only look at a small group of games. This paper examines the dynamics on an entire class of games. These games will have two Pareto rankable equilibria, but no dominant strategies.

There has also been some work which examines a larger class of games. Rankin, Van Huyck, and Battalio (2000) run experiments with 75 periods of stag-hunt games where they vary the payoff value, $x \in (0, 1)$, uniformly with an error term of ε_t each period. They find that the payoff dominant equilibrium is usually selected even with high values of x which make it a risky decision. Stahl and Van Huyck (2002) perform similar experiments but consider two different classes of games. They induce coordination by showing subjects games in which coordination is easy, and this carries over to games where coordination is difficult. They analyze the effect of the previous state on the current decision. Both of these papers are considering a wide class of games. However, they perform random dynamics over the class of games. This makes it more difficult to identify certain bifurcations that may exist in the equilibrium correspondence. In this paper, the dynamics are organized by slightly changing one parameter at a time in an organized manner which allows easier identification of the bifurcations and hence the hysteresis.

There have also been some experimental papers which provide evidence of this type of hysteresis in coordination games. Weber (2006) examines the effect of changing group sizes on the group's ability to coordinate in a minimum-effort coordination game. He finds that coordination in large groups is possible if the group starts with a small number of subjects, and gradually increases to a size of 12 subjects per group. This is in contrast to groups that start with 12 subjects per group, which are never able to coordinate on high-effort levels. This suggests that there is hysteresis based on the group size, because the selected equilibrium for group size 12 depends on the history leading up to that game. Brandts and Cooper (2006) examine the effect of using payoff bonuses as means of inducing cooperation in the minimum-effort coordination game. They find that adding bonuses helps bring groups from low effort levels to higher effort levels. In addition, they find that when the effort levels are decreased back to initial levels after the temporary bonuses, effort levels are higher than before the bonuses. This dependence on the history of the game is the type of behavior that I study in this paper.

Another example of a coordination game is a platform competition game, such as which operating system an individual should use. If all of your collaborators use operating system A, then it makes sense for you to use operating system A, because this allows for easier collaboration. Similarly, if all of your collaborators use operating system B, it is beneficial for you to also use B. In this situation, it is beneficial for people to coordinate with each other on the same platform. An example of this is the QWERTY keyboard. David (1985) suggests that the QWERTY-keyboard is suboptimal when compared to the DVORAK arrangement. However, due to the headstart by QWERTY, it has prevailed as the main type of keyboard used today.

Why should we expect to see hysteresis in a coordination game like this? Continuing with the keyboard example, suppose that people can choose either the QWERTY or the DVORAK keyboard. At time $t = 0$, QWERTY (100WPM) is determined to be better than DVORAK (50WPM), and so all people start to use this keyboard. At some later time, $t = 1$, suppose the language has changed, and now QWERTY (70WPM) is no longer better than DVORAK (80WPM). Even though QWERTY is no longer optimal, people have become accustomed to this arrangement, and therefore continue to use it. At time $t = 2$, the language has changed more so QWERTY (50WPM) is even more suboptimal than DVORAK (100WPM). The convenience of using QWERTY is now outweighed by the cost of using a suboptimal keyboard, so people start to switch to the DVORAK layout. At time $t = 3$, the language has changed back, so that QWERTY (70WPM) is almost as good as DVORAK (80WPM) again. Now, people are accustomed to the DVORAK layout and since it is optimal they continue using DVORAK. So even though the situation was the same at $t = 1$ and $t = 3$ (QWERTY

70WPM vs. DVORAK 80WPM), a different keyboard is selected at $t = 1$ than $t = 3$. This hysteresis is caused by the fact once the players coordinate on a certain outcome, it may be difficult to change because players do not want to incur the switching costs. I find that the same type of hysteresis occurs in coordination games, and this intuition is formalized in this paper.

4.1.1 Bifurcations

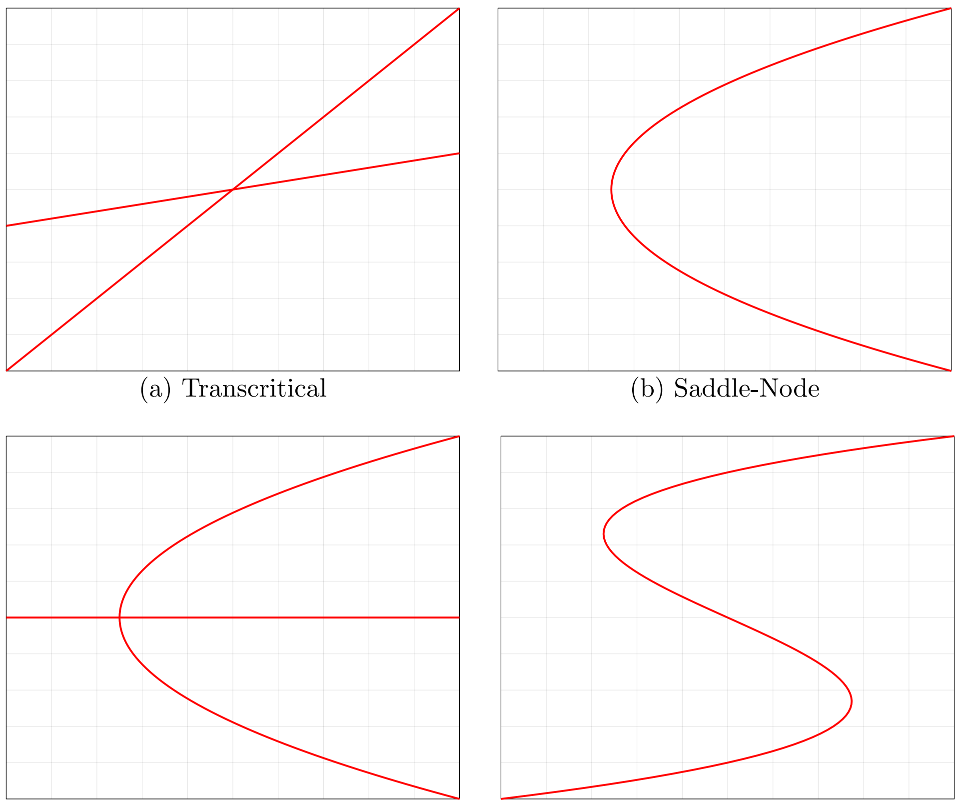
Whenever there are multiple equilibria, it is likely that there exist bifurcations of equilibrium points into multiple equilibria as a parameter is varied. Understanding these bifurcations may help determine which equilibrium outcome is more likely for any given situation. Even though deriving analytical solutions for these bifurcation points is often difficult, equilibrium correspondences based on certain parameters of the system can be plotted using numerical methods. In many systems, as the parameter is varied, the correspondence may bifurcate, meaning solutions can appear and disappear at critical points. Before getting in to the model, it is important to introduce several types of bifurcations.

Transcritical - A transcritical bifurcation is one such that two solutions cross each other. This can be seen in Figure 4.1(a). When the two solution cross, they change stability. That is, if the lower one is stable before the intersection, after they cross, this solution is now the higher solution, and unstable.

Saddle-Node A saddle-node bifurcation is a one such that there are zero solutions to one side of the critical point, two solutions on the other side, and exactly one at the critical bifurcation point. A saddle-node bifurcation with critical point $x = 0$ is displayed in Figure 4.1(b).

Pitchfork A pitchfork bifurcation has one solution to one side of the critical point and three on the other side of the critical point. A pitchfork bifurcation at $x = 0$ is shown in Figure 4.1(c).

Any combination of these bifurcations can be present in the bifurcation diagram for a given model. One important combination is the double saddle-node bifurcation. A double saddle-node is a bifurcation such that there is a unique solution in the limit of each direction, but three solutions



(a) Transcritical (b) Saddle-Node (c) Pitchfork (d) Double Saddle-Node

Figure 4.1: Types of Bifurcations.

inside some region in the middle. The double saddle-node bifurcation takes the form of an s-shaped curve, and is shown in Figure 4.1(d). For very low values of x, there is one solution, but then the saddle-node bifurcation at $x = -0.4$ increases the number of solutions from one to three. Further, the saddle-node bifurcation at $x = 0.4$ decreases the number of solution back from three to one. This type of bifurcation is the main focus of this paper.

If solutions tend to stay on the solution path (they do not jump between the equilibrium solutions), then the double saddle-node bifurcations leads to hysteresis. For example, consider Figure 4.2. This shows a double saddle-node bifurcation. In this example, suppose that the parameter γ varies first from 0.2 to 0.8 (denoted by the red line), then it decreases from 0.8 back to 0.2 (denoted by the blue line). In this case, the system starts at $\gamma = 0.2$, where there is a unique equilibrium. As γ increases, the system remains on the top part of the s-shaped curve, until it reaches the saddle-node bifurcation at $\gamma = 0.58$ at which point the high equilibrium ceases to exist. For $\gamma > 0.58$, there is a unique equilibrium so the system will jump from the high equilibrium down to the low equilibrium (jump denoted by dotted line). The system remains in the low equilibrium as γ increases from 0.58 to 0.8. When γ decreases from 0.8, it remains on the low solution until $\gamma = 0.42$, at which point the low equilibrium ceases to exist, and again there is a unique equilibrium. This causes the system to jump back up from the low equilibrium to the high equilibrium. Therefore, for intermediate values, $\gamma \in (0.42, 0.58)$, the outcome depends on the starting position. When starting with $\gamma = 0.2$, the system goes to the equilibrium $\sigma = 0.94$ at $\gamma = 0.5$. When starting with $\gamma = 0.8$, the system goes to the equilibrium $\sigma = 0.058$ at $\gamma = 0.5$. This hysteresis is caused by the s-shaped equilibrium correspondence, and the assumption that the system traces along the equilibrium correspondence as the parameter changes.

The next section introduces the model, and then examines situations where hysteresis occurs.

4.2 Model

Let $g(\gamma)$ be a game consisting of n players, $I = \{1, 2, \ldots, n\}$. Each player has m pure actions, $S_i = \{s_i^1, s_i^2, \ldots, s_i^m\}$. A joint-action profile is denoted by $\boldsymbol{s} = \{s_1, \ldots, s_n\}$. Each player faces a payoff

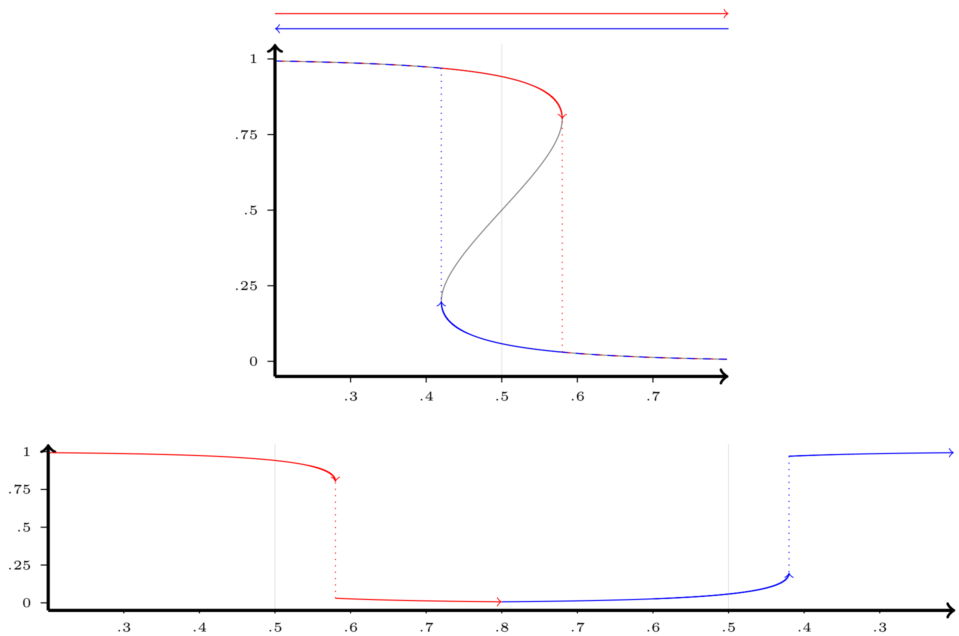


Figure 4.2: Example of how double saddle-node bifurcation leads to hysteresis.

function $u_i(\boldsymbol{s}, \gamma)$ which depends on the parameter γ from parameter space Γ. Let $\mathcal{G} = \{g(\gamma) | \gamma \in \Gamma\}$ be the set of all games over parameter space Γ.

The set of mixed strategies is denoted by $\Sigma_i = \Delta^i$, which is the set of probability distributions over S_i. A mixed strategy is denoted by $\sigma_i \in \Sigma_i$, which is a mapping from S_i to Σ_i, where $\sigma_i(s_j)$ is the probability that player i plays pure-action s_j, and $\Sigma = \Sigma_1 \times \cdots \Sigma_n$ is the set of mixed strategy profiles. A joint mixed-strategy profile is denoted $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_n\}$. Player i's expected payoff for mixed-strategy profile $\boldsymbol{\sigma}$ is $u_i(\boldsymbol{\sigma}, \gamma) = \sum_{\boldsymbol{s} \in S_1 \times \cdots \times S_n} p(\boldsymbol{s}) u_i(\boldsymbol{s}, \gamma)$, where $p(\boldsymbol{s}) = \Pi_{i \in I} \sigma_i(s_i)$ is the probability of the pure-strategy profile $\boldsymbol{s}$ give mixed strategy profile $\boldsymbol{\sigma}$.

A joint strategy profile $\boldsymbol{\sigma}$ is an equilibrium of the game $g(\gamma)$ if the equilibrium function $f : \Sigma \times \Gamma \times \Lambda \to \mathbb{R}$, dependent on parameter $\lambda \in \Lambda$ satisfies $f(\boldsymbol{\sigma}, \gamma, \lambda) = 0$. For example, f could be the logit quantal response equilibrium function,

$$f(\boldsymbol{\sigma}, \gamma, \lambda) = \sum_{i=1}^{m} \left| \frac{e^{\lambda u_l(s_i, \boldsymbol{\sigma}_{-i}, \gamma)}}{\sum_{j=1}^{m} e^{\lambda u_l(s_j, \boldsymbol{\sigma}_{-i}, \gamma)}} - \sigma_L(s_i) \right| = 0 \tag{4.1}$$

Given γ and λ, any joint mixed-strategy profile $\boldsymbol{\sigma}$ is an equilibrium if (4.1) is satisfied. The game

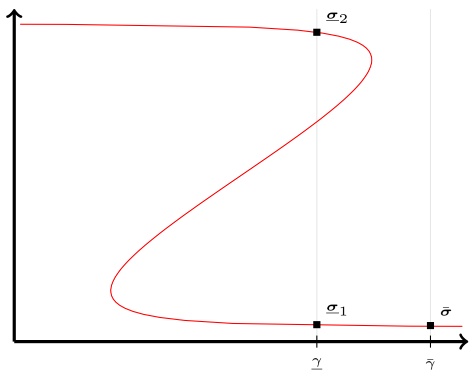


Figure 4.3: Example of continuous variation of a correspondence between two parameter values.

$g(\gamma)$ has multiple equilibria if $f(\boldsymbol{\sigma}, \gamma, \lambda) = 0$ for more than one joint mixed-strategy profile $\boldsymbol{\sigma}$. Let $\boldsymbol{\Sigma}^*(\gamma, \lambda) = \{\boldsymbol{\sigma} | f(\boldsymbol{\sigma}, \gamma, \lambda) = 0\}$ be the set of equilibria of game $g(\gamma)$ according to equilibrium function f with parameter λ.

Definition 4.2.1. *The equilibrium correspondence* $\boldsymbol{\Sigma}^*(\gamma, \lambda)$ *varies continuously from* $\underline{\gamma}$ *and* $\bar{\gamma}$ *starting at* $\underline{\boldsymbol{\sigma}} \in \boldsymbol{\Sigma}^*(\underline{\gamma}, \lambda)$ *if and only if for all* $\varepsilon > 0$*, there exists some* $N \in \mathbb{N}$ *such that for all* $k = 0, \ldots, N-1$*,*

$$\gamma_{k+1} - \gamma_k = \frac{\bar{\gamma} - \underline{\gamma}}{N} \Rightarrow \|\boldsymbol{\sigma}_{k+1} - \boldsymbol{\sigma}_k\| < \varepsilon$$

where $\underline{\gamma} = \gamma_0, \gamma_N = \bar{\gamma}$ *and* $\boldsymbol{\sigma}_k \in \boldsymbol{\Sigma}^*(\gamma_k, \lambda)$*. The endpoint of this continuous path is* $\bar{\boldsymbol{\sigma}} \in \boldsymbol{\Sigma}^*(\bar{\gamma}, \lambda)$*.*

The equilibrium correspondence varies continuously if you can trace the correspondence between the two parameter values while always moving in the direction from the first parameter value to the second parameter value. For example, in Figure 4.3, the equilibrium correspondence varies continuously from $\underline{\gamma}$ to $\bar{\gamma}$ starting at $\underline{\boldsymbol{\sigma}}_1$. However, the equilibrium correspondence *does not* vary continuously from $\underline{\gamma}$ to $\bar{\gamma}$ starting at $\underline{\boldsymbol{\sigma}}_2$.

Assumption #1: When faced with $g(\gamma)$, players will play one of the equilibria, call this $\boldsymbol{\sigma}(g(\gamma)) \in \boldsymbol{\Sigma}^*(\gamma, \lambda)$.

Assumption #2: If game $g(\gamma)$ is played, and players play equilibrium $\boldsymbol{\sigma}(g(\gamma))$, then when game $g(\gamma')$ is played, if the equilibrium correspondence $\boldsymbol{\Sigma}^*(\gamma, \lambda)$ varies continuously from γ to γ' starting at $\boldsymbol{\sigma}(g(\gamma))$ with endpoint $\boldsymbol{\sigma}'$. Then when $g(\gamma')$ is played, players will play $\boldsymbol{\sigma}'$, that is

$\boldsymbol{\sigma}\left(g\left(\gamma'\right)\right) = \boldsymbol{\sigma'} \in \boldsymbol{\Sigma}^*\left(\gamma', \lambda\right)$.

Definition 4.2.2 (Hysteresis)**.** *The equilibrium correspondence $\boldsymbol{\Sigma}^*\left(\gamma, \lambda\right)$ exhibits hysteresis for equilibrium function f with parameter λ if there exists points $\gamma_1, \gamma_2, \gamma_3$ such that,*

1. *the correspondence varies continuously from γ_1 to γ_2 with starting point $\boldsymbol{\sigma}_1 \in \boldsymbol{\Sigma}^*\left(\gamma_1, \lambda\right)$ and endpoint $\boldsymbol{\sigma}_2 \in \boldsymbol{\Sigma}^*\left(\gamma_2, \lambda\right)$,*
2. *the correspondence varies continuously from γ_3 to γ_2 with starting point $\boldsymbol{\sigma}_3 \in \boldsymbol{\Sigma}^*\left(\gamma_3, \lambda\right)$ and endpoint $\boldsymbol{\sigma}'_2 \in \boldsymbol{\Sigma}^*\left(\gamma_2, \lambda\right)$, and*
3. *$\boldsymbol{\sigma}_2 \neq \boldsymbol{\sigma}'_2$.*

What types of games have this hysteresis property? The next section examines the minimum-effort coordination game, and shows that it exhibits hysteresis.

4.2.1 Minimum-Effort Coordination Game

A minimum-effort coordination game consists of n players, $I = \{1, \ldots, n\}$. Each player has two actions, they can either choose to exert high effort or low effort, $S_i = \{x_L, x_H\}$ for $x_L, x_H \in \mathbb{R}$ and $x_L < x_H$. The joint pure-action profile is denoted by $\boldsymbol{s} \in \{x_L, x_H\}^n$. Performing the high effort is more costly than performing the low effort. The benefit of the high effort is only received if every player plays the high effort action. If any player chooses the low effort action, then all players only receive the benefit from the low action. This yields payoffs,

$$u_i\left(\boldsymbol{s}\right) = \min_{j=1,\ldots,n} s_j - cs_i.$$

The normal form of the minimum-effort coordination game is displayed in Figure 4.4. For a given value of cost, $c \in \mathbb{R}$, the minimum-effort coordination game is denoted by $cg(c)$. The set of all minimum-effort coordination games is $\mathcal{CG} = \{cg(c) | c \in \mathbb{R}\}$.

Player i	$s_j = x_h$ for all $j \neq i$	$s_j = x_L$ for some $j \neq i$
x_H	$x_H (1 - c)$	$x_L - cx_H$
x_L	$x_L (1 - c)$	$x_L (1 - c)$

Figure 4.4: Minimum-Effort Coordination Game.

4.2.2 Nash Equilibria

If $c > 1$, then the cost of exerting high effort outweighs the benefit, so action x_L strictly dominates x_H for all players. Therefore all players playing x_L is the pure-strategy Nash equilibrium when $c > 1$. Similarly, if $c < 0$, then the cost is negative, so the action x_H strictly dominates x_L. So, all players playing x_H is the unique Nash equilibrium. This game is most interesting when there are multiple equilibria, which occurs when $c \in [0, 1]$. For these intermediate cost values, the game has two pure strategy Nash Equilibria: one where everyone plays the high effort x_H, and one where everyone plays the low effort x_L. There is also one symmetric mixed strategy equilibrium where all players play x_H with probability $c^{\frac{1}{N-1}}$, which is clearly increasing in N for $c \in (0, 1)$.

For all values of $c \in (0, 1)$, the equilibrium where all players play x_H with probability 1 is the *payoff dominant* equilibrium. For levels of c close to 1, the difference between the high effort equilibrium payoff and the low effort equilibrium payoff becomes small. However, there is a large loss possible if the high effort action is played, while there is no loss possible if the low effort is played. Therefore, when c is close to 1, the high effort action is risky. In fact, when $c < {}^1\!/_2{}^{N-1}$, everyone playing the high effort is the *risk dominant* equilibrium. When $c > {}^1\!/_2{}^{N-1}$, everyone playing the low effort is the risk dominant equilibrium.

4.2.3 Symmetric Quantal-Response Equilibria

This section studies properties of the symmetric quantal-response equilibria. Suppose that $\boldsymbol{\sigma}_H^*$ is the probability that a player plays the high effort action x_H, and $\boldsymbol{\sigma}_L^* = 1 - \boldsymbol{\sigma}_H^*$ is the probability that a player plays the low effort equilibrium. Using the logit quantal response function, a SQRE must satisfy the equation,

$$\begin{aligned} \boldsymbol{\sigma}_H^* &= \frac{e^{\lambda u_i(H,\boldsymbol{\sigma}_{-i}^*)}}{e^{\lambda u_i(L,\boldsymbol{\sigma}_{-i}^*)} + e^{\lambda u_i(H,\boldsymbol{\sigma}_{-i}^*)}} \\ &= \frac{1}{1 + e^{\lambda\left[u_i(H,\boldsymbol{\sigma}_{-i}^*) - u_i(L,\boldsymbol{\sigma}_{-i}^*)\right]}} \\ &= \frac{1}{1 + e^{\lambda\left[(x_h - x_L)\left(\boldsymbol{\sigma}_H^{*\,N-1} - c\right)\right]}}. \end{aligned} \tag{4.2}$$

This nonlinear equation has two parameters to vary: c and N. There are a couple of important points to consider. When $\lambda = 0$, there is always a unique SQRE, $\boldsymbol{\sigma}_H^* = 0.5$. The intuition for this, is that when one player is playing randomly ($\lambda = 0$), then it is the best response for the other players to play randomly as well. Secondly, in the limit as $\lambda \to \infty$, there are always three solutions because the set of SQRE approaches the set of Nash equilibria as $\lambda \to \infty$, and there are always three symmetric Nash Equilibrium for games with $c \in (0, 1)$. Therefore, since there is one SQRE at $\lambda = 0$ and three at $\lambda = \infty$, there must be a bifurcation at at least one value of $\lambda \in (0, \infty)$. Due to the non-linearity of (4.2), an analytical solution is not feasible and therefore it must be solved numerically. Finding these solutions requires using Newton's method, as well as being careful not to overlook extra solutions. Most of these solutions were obtained by doing multiple swipes of the parameter space to ensure that all solutions were found.

4.2.3.1 Varying λ

First, I examine the effect of changes in the λ parameter on the SQRE correspondence. As mentioned above, for low values of λ there is a unique solution, and for high values of λ there are three equilibrium values of $\boldsymbol{\sigma}_H^*$, so the correspondence is likely be a pitchfork bifurcation or something similar.

Proposition 4.2.3. *Using the symmetric quantal response equilibrium function f with parameter λ, the equilibrium correspondence $\mathbf{\Sigma}^* (c, \lambda)$ as λ varies has the following properties:*

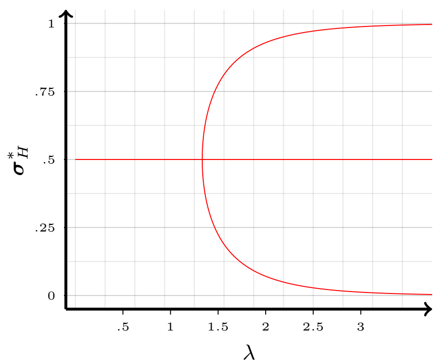
1. *$1/2 \in \mathbf{\Sigma}^* \left(\frac{1}{2^{N-1}}, \lambda\right)$ for all λ.*

2. *The correspondence varies continuously from $\lambda = 0$ to $\lambda = \infty$ with endpoint $\boldsymbol{\sigma}_H = 1$ if $c < \frac{1}{2^{N-1}}$ and endpoint $\boldsymbol{\sigma}_H = 0$ if $c > \frac{1}{2^{N-1}}$.*

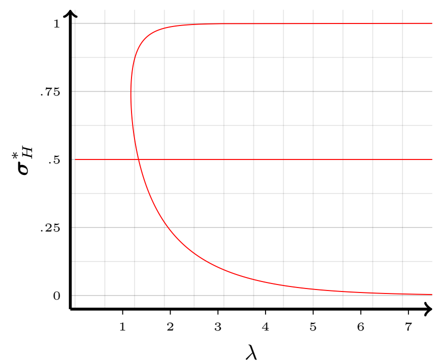
The SQRE correspondences for different values of N are displayed in Figure 4.5. These are all graphed for the cost $c = 1/2^{N-1}$, because that cost ensures that $\boldsymbol{\sigma}_H^* = 0.5$ is a solution for all values of λ. These are all similar looking to a pitchfork bifurcation because the choice of cost. However, the only pitchfork bifurcation is $N = 2$. The others are a combination of first a saddle-node bifurcation, and then a transcritical bifurcation. For example, in the $N = 10$ case, there is one solution at $\lambda = 0$, then there is a saddle-node bifurcation at about $\lambda = 1$, where the number of solutions changes from one to three. Finally at $\lambda = 40$ there is a transcritical bifurcation where two of the solutions cross paths.

The intuition for these graphs is quite straightforward. If $\lambda = 0$ then there is no response to the payoffs, so the players are playing randomly. Since this is a coordination game, if the other players are playing randomly, then the best response is to play randomly as well. As N gets larger, the graph becomes compressed away from the λ-axis, meaning that x_L, the low-effort action is played with lower probability in equilibrium. When there are a lot of players, playing the high-effort action is risky, and therefore in equilibrium player i only plays x_H with high probability if the other players are also playing x_H with high probability, even for low values of λ. As N increases this risk becomes greater, meaning that the the equilibrium probability of playing x_H, must be increasing in N.

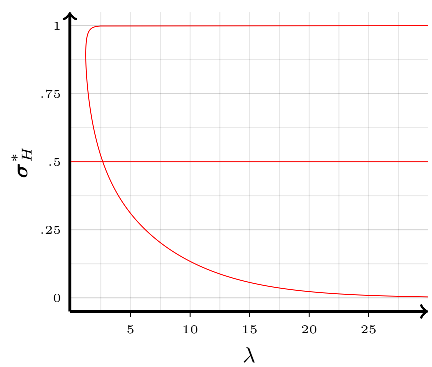
The second part of Proposition 4.2.3 shows that $c = \frac{1}{2^{N-1}}$ is a critical value for the SQRE correspondences. When $c < \frac{1}{2^{N-1}}$, the correspondence varies continuously from $\lambda = 0$ to $\lambda = \infty$ with endpoint $\boldsymbol{\sigma}_H^* = 1$. An example of this is shown in Figure 4.6. The first panel in the figure shows the correspondence for $c = 0.495$. The pitchfork bifurcation breaks into two parts: one part



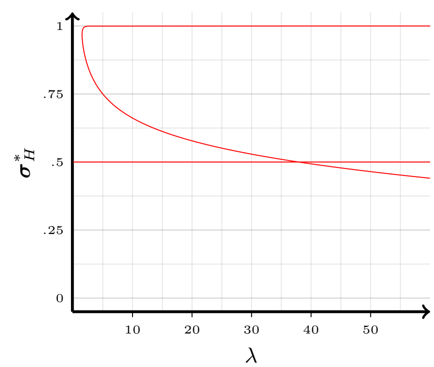
(a) $N = 2$



(b) $N = 3$

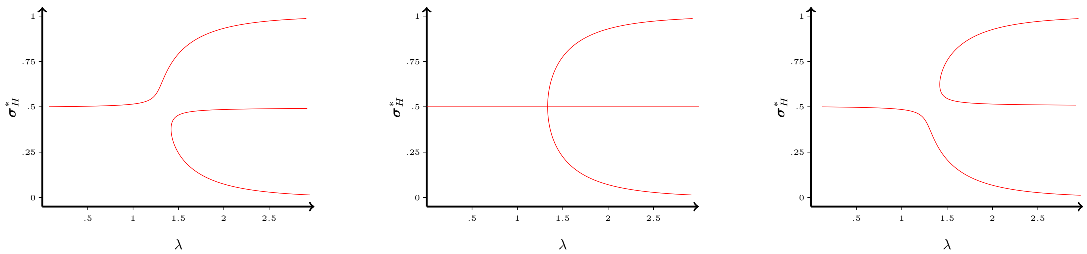


(a) $N = 5$



(b) $N = 10$

Figure 4.5: QRE correspondences as λ is varied for different value of N and $c = {}^1\!/_2{}^{N-1}$.



$c = 0.495$ $c = 0.5$ $c = 0.505$

Figure 4.6: Example of how pitchfork bifurcation breaks apart.

increasing continuous from $\lambda = 0$ to $\lambda = \infty$ and another turning point around $\lambda = 1.5$. Similarly when $c > \frac{1}{2^{N-1}}$, the pitchfork also breaks into two parts. The first part is again a continuous from $\lambda = 0$ to $\lambda = \infty$, however it is now decreasing. There is also a turning point around $\lambda = 1.5$. The value $c^* = \frac{1}{2^{N-1}}$ is important because that is the critical value for risk dominance. When $c > c^*$, the risk dominant equilibrium is for all players to choose x_L, and when $c < c^*$, the risk dominant equilibrium is for all players to play x_H. So the risk dominant equilibrium can be determined by following the correspondence path from $\lambda = 0$ to $\lambda = \infty$. However, the other equilibria appear for higher values of λ. In the next section, I examine the equilibrium correspondence for variations in the cost parameter c.

4.2.3.2 Varying c

If an equilibrium correspondence has a pitchfork bifurcation at parameter value λ^*, then it is often the case that for fixed $\lambda > \lambda^*$, variation of another parameter leads to a double saddle-node bifurcation. To better understand why this is true, it helps to look at a three-dimensional representation of the parameter space when there is a pitchfork bifurcation as displayed in Figure 4.7.

The previous section shows that when $N = 2$ and $c = 0.5$, the equilibrium correspondence has a pitchfork bifurcation at $\lambda = 1.35$. Therefore, for fixed $\lambda > 1.35$, a double saddle-node bifurcation is likely to occur when another parameter is varied. This turns out to be the case in the symmetric quantal-response equilibrium correspondence of the minimum-effort coordination game as seen in the following proposition,

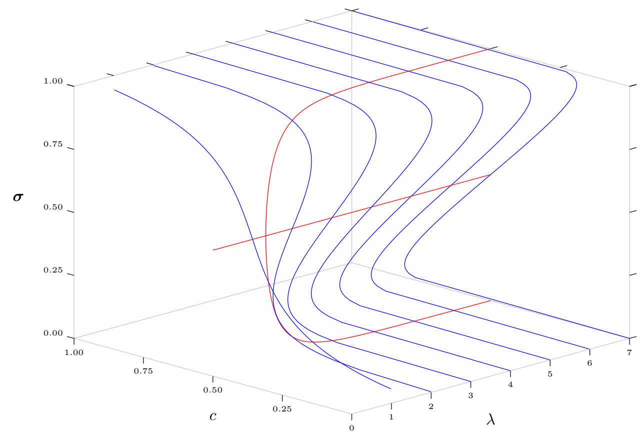


Figure 4.7: Example showing the relationship between a pitchfork bifurcation and a double saddle-node bifurcation.

Proposition 4.2.4. *For every coordination game, $g(c) \in \mathcal{CG}$:*

1. *There exists a λ^* such that the logit SQRE correspondence, $\mathbf{\Sigma}^*(c,\lambda)$, exhibits hysteresis for all $\lambda > \lambda^*$, where,*

$$\lambda^* = \left(\frac{N}{N-1}\right)^N \frac{1}{x_h - x_l}.$$

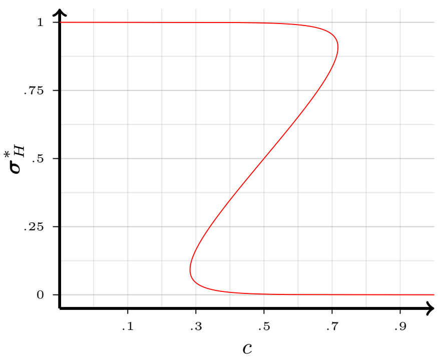
2. *The critical value λ^* is decreasing in N.*

3. *For $N = 2$, the saddle-node bifurcation points are given by,*

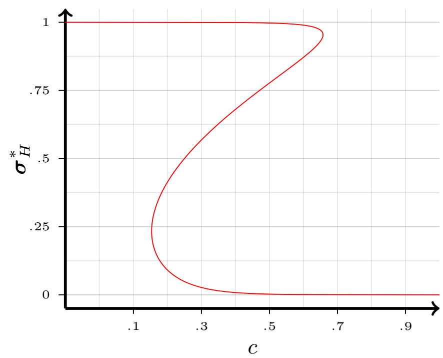
$$\sigma_H = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4}{\lambda\,(x_h - x_L)}}.$$

4. *For $N > 2$, the saddle-node bifurcation points will not be symmetric around $\sigma_H = \frac{1}{2}$.*

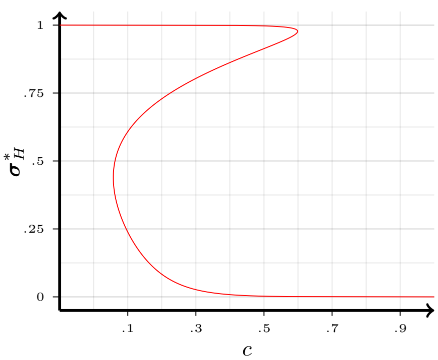
This proposition states that for any coordination game of the given form, if the players have a high enough payoff responsiveness (sufficiently high λ), then the game should exhibit hysteresis. This means that given a game and an equilibrium, it is possible to vary one parameter slightly and then change it back, and the system could be at a completely different equilibrium. This can be very important if one of the equilibria is more desirable then the other and all that is required is a small perturbation of the system to go from the less desired to the more desired equilibrium. The



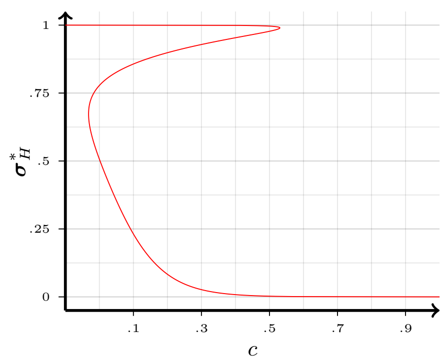
(a) $N = 2$



(b) $N = 3$



(a) $N = 5$



(b) $N = 10$

Figure 4.8: QRE correspondences as c is varied for different value of N and $\lambda = 4$.

second point in the proposition says that this critical value is decreasing as the size of the group gets larger. Assuming that the values of λ for the individuals are not dependent on group size, then this means that hystersis is more likely as the group size increases. The third part gives the analytical solution for the values of the saddle-node bifurcations for the $N = 2$ case. It is not possible to find the analytical solution for the $N > 2$ case, but with numerical analysis, it is clear that the c value of both saddle-nodes is decreasing as the group size gets larger (as shown in Figure 4.8). Also, for group size larger than two, the double saddle-node bifurcation is not be symmetric, meaning that the bifurcation points are not equidistant from 0.5. However, in the $N = 2$ case, the double saddle-node bifurcation is symmetric.

The proof of this is given in the appendix. The proof of this proposition involves analyzing certain properties of the equilibrium correspondence. The double saddle-node bifurcation is not a function, so it is difficult to analyze. However, it is possible to solve for the equilibrium value of c as a function of σ_H, $c^*(\sigma_H)$, which is a function and therefore easier to work with. The equilibrium correspondence $\mathbf{\Sigma}^*(c, \lambda)$ has a double saddle-node bifurcation if $c^*(\sigma_H) \to \infty$ as $\sigma_H \to 1$, $c^*(\sigma_H) \to -\infty$ as $\sigma_H \to 0$, and $c^{*\prime}(\sigma_H) < 0$ for some value of $\sigma_H \in (0, 1)$. This process ensures existence of a double saddle-node bifurcation for sufficiently large values of λ. The rest of the proposition is obtained from comparative statics which are detailed in the appendix.

This setup yields some testable implications, the most important being that the equilibrium correspondence of the minimum-effort coordination game exhibits hysteresis. To test this, it is necessary to run an experiment with multiple games, where the games are varied in an organized manner to determine whether the experimental outcomes exhibit hysteresis or not. In the next section I detail the experimental methods and results.

4.3 Experiments

4.3.1 Experimental Design

There were five sessions for this experiment, each consisting of between 26-34 rounds. The subjects were drawn from a pool of both graduate and undergraduate students from the California Institute of Technology that have signed up for the Social Sciences Experimental Laboratory.

The subjects were seated at computers, and read instructions as they were displayed in a slide show. All questions were answered in front of everybody, to ensure that all the information was common knowledge. Each subject played the simple game described above with two actions. Each round the subjects were randomly divided up into a group, and these groups changed every round (except for the fifth treatment in which case there was only one group). Each round they choose between two actions, X and Y, and for any given round, every player in the group faced the same decision.

Over the course of the experiment the cost parameter was changed in an ordered matter. Some treatments it was changed from low to high to low (LHL), in others it was changed from high to low to high (HLH), and in the first treatment the experiment started at the $c = 0.5$ game (MLHL) as a baseline to see how subjects acted initially. One nice feature about this experiment is that if the cost is very low ($c < 0$), then playing the high-effort action strictly dominates playing the low-effort action. Also if the cost is very high ($c > 1$), then playing the low-effort action strictly dominates playing the high-effort action. Therefore, if the cost is varied from a cost below zero to a cost above one, and if the players are not playing strictly dominated strategies, then they should pick each action at least once.

The payoff in each round was determined by both the individual's action, and the actions of the other members of the individual's group (which ranged from 2-12 in the different treatments). The subjects' overall payoff was the sum of each individual round payoff. The exchange rate ranged from 30 to 40 cents per point. The players received one point if some player in their group chose the low-effort action, and received four if all players in their group chose the high-effort action. In

	Session 1	Session 2	Session 3	Session 4	Session 5
Group Size	2	2	5	10	12
Number of Groups	6	8	3	2	1
Number of Rounds	26	28	30	33	34
Order	MLHL	HLH	HLH	HLH	LHL
Cost Range	$.1 - .9$	$0 - 1$	$0 - 1$	$-0.05 - 1.05$	$-0.05 - 1.05$
Matching	Random	Random	Random	Random	Only 1 Group
Actions	X-Low/Y-High	Random	Random	Random	Random
Exchange Rate	\$0.40	\$0.30	\$0.30	\$0.30	\$0.40

Table 4.1: A summary of the experimental sessions.

addition, the players incurred a cost of c if they played the low-effort action, and a cost of $4c$ if the played the high-effort action.

The subjects were given information on the actions of the other group members through both their own payoffs and the payoffs of the other members of the group. Based on this information they could infer the actions of all other members of their group. In addition the history for each of the previous rounds was available over the course of the entire experiment. This history contained their action, their payoff, and the payoffs of all other group members. A summary of the different sessions is given in Table 4.1. The subjects were asked to record the outcome of the previous round before moving on to the next round, in order to make sure that they were paying attention. The record sheet consisted of the round number, their action, their group's action, and their payoff for that round.

We decided not to do a practice round before the experiment for two reasons. First the game is quite simple, and not much could be learned from an example. To make sure that the subjects were comfortable with the game, the instructions contained two detailed examples, and gave an overview of the experimental interface. The second reason for having no practice round is that it could have a large effect on the groups ability to coordinate. In practice rounds, since the subjects are not getting paid, they will sometimes just choose randomly, or choose the wrong action on purpose. This could give subjects a false impression of other members of their group. In addition the subjects could also use it to their advantage in order to get a head start on organizing coordination. For these reasons, no practice round was given. Even though there was no practice round, all of the sessions, with the exception of the first session, started with a game that had either a strictly or weakly dominant strategy.

4.3.2 Experimental Results

In the first session, there were 12 subjects total, divided into groups of 2. The subjects first played three rounds with $c = 0.5$. These three rounds served as a proxy for how the subjects played without any prior experience. However, these preliminary rounds were dropped in the later sessions, because they could be used as a coordination device, or inhibit the group's ability to coordinate. The first session proved that Caltech students are able to coordinate well, especially when the experiment starts at a parameter value that promotes coordination. In fact, the subjects only chose the low-effort action 6 out of 312 choices. Even though the game started at a medium risk level of $c = 0.5$, they were immediately able to coordinate, which was surprising based on the previous literature.

Another problem with this experiment was that the subjects had two actions, X always corresponded to the low-effort action, and Y always corresponded to the high-effort action. In the middle of the experiment, it was noted that some of the subjects had already filled out their choice for all of the remaining rounds of the experiment. This means that they were not taking the changing costs into consideration. In order to fix this problem, the action labels were randomized before the experiment using a random number generator in MATLAB.

The second session had 16 subjects that were paired into groups of two, and played 28 rounds. In the first 5 rounds, the subjects faced the game where $c = 1$. In this game, they had a choice between a guaranteed payment of 0 or, a chance between 0 and -3 depending on what their partner played. In this game, the high action was a weakly dominated strategy and the prediction was that the weakly dominated strategy would never be played. However, in the experiment, they choose the weakly dominated strategy 15% of the time (12 out of 80, but nobody picked it more than twice). By the fifth round of $c = 1$ the subjects all played the low-effort action, as predicted.

In the sixth round, the cost was decreased to $c = 0.9$ at which point 2 out of the 16 subjects played the low-effort action, while all the others played the high-effort action. This was surprising, because the $c = 0.9$ game is quite risky. The subjects could either guarantee 0.1 by playing the low-effort action, or take a chance between 0.4 and -2.6 by playing the high-effort action. These two subjects out of 16 were not enough to get more subjects to play the low-effort action. Therefore for the next 22 rounds, every single subject played the high-effort action without a single deviation. Again, this level of coordination was unprecedented based on the previous literature.

Based on the success of the subjects in the first session, the next sessions were run with a larger group size, in which the subjects would not be able to coordinate as easily. Based on the second point of Proposition 4.2.4, this hysteresis is more likely with larger groups. Session three was the first large group session with 15 subjects that were grouped into 3 groups of 5 subjects each. The subjects started with three rounds of the $c = 1$ game, where playing the low-effort action is weakly dominated. Surprisingly, after changing to the $c = 0.9$ game in the third round, only 3 out of 15 played the low-effort action. Quite quickly, the subjects were all able to coordinate on the high-effort action, even when they started by coordinating on the low-effort action. Then after the initial few who played the low-effort action, they were able to coordinate almost perfectly for the remainder of the session. The experimental data is shown in Figure 4.9(b). One subject played the low-effort action from the beginning until the cost reached $c = 0.5$, but then played the high-effort action on the way back. This individual played differently in the first $c = 0.5$ game than in the second $c = 0.5$ game. This is the type of hysteresis that was predicted to happen at the group level. With the exception of this subject, the group was able to coordinate too well to show the hysteresis that was hypothesized.

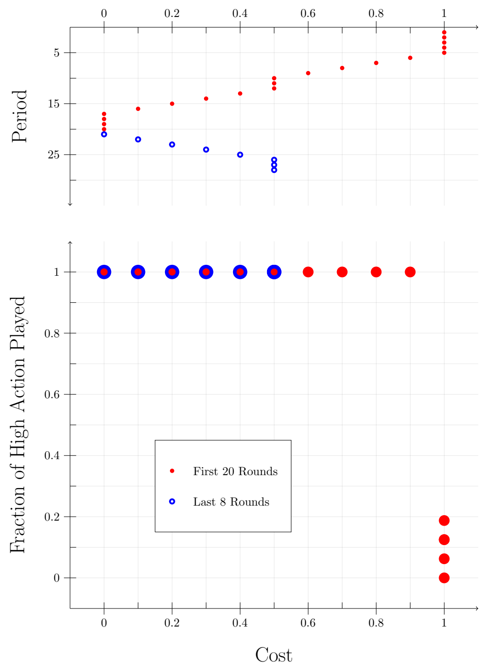
The first three sessions all gave surprising results. Typically subjects have trouble coordinating in minimum-effort games, especially with group sizes larger than two. However even in the third session with group size five, the Caltech subjects were able to coordinate almost perfectly. The reason for this could be that Caltech students participate in a lot of experiments together, and this helps them to coordinate better than most subjects.

In order to show the desired result of hysteresis, the group size was increased even more than the first three sessions. The fourth session contained 20 subjects with two groups of 10. This session started with the high cost $c = 1$, in which playing the low-effort action is a strictly dominant strategy. Subjects were able to play the strictly dominant strategy 39 out of 40 times in the first two rounds. In the third round, five of the subjects tried to get the group to coordinate by playing the high-effort action, but this did not work, so most of the group went with the low-effort, secure action. As the cost decreased, the subjects remained playing the low-effort action except for one subject. One subject continued to play the high-effort action throughout the first half of the experiment. By the time the experiment got to the $c = 0$ game, this subject was at -\$10, at which point he decided to "spite" the rest of the group.[1] So in the $c = 0$ game, this subject was the only one to play the weakly dominated strategy, and in the $c = -0.05$ game this subject was the only one to play the strictly dominated strategy. The other subjects in the groups noted this irrational behavior, and therefore were not able to coordinate as c increased. The experimental data for session 4 is given in Figure 4.9(c). This individuals behavior had a large impact on the group's decisions, and therefore made group coordination very difficult.

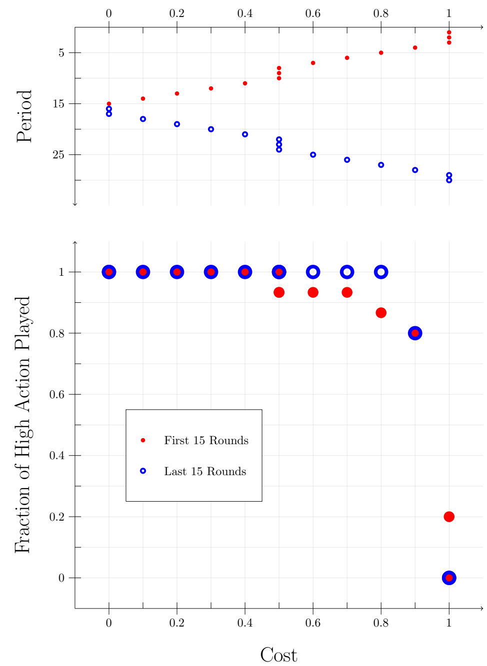
One hypothesis for this may be that this hysteresis is very fragile, and one tremble can have an immediate impact on the ability to coordinate. But this was more than just one tremble, it was back-to-back rounds where an individual played a dominant strategy. Even under the QRE model, which allows for trembles, the probability of a tremble at the $c = -.05$ for $\lambda = 4$ is 3.3732×10^{-6}. So even one individual playing the strictly dominated strategy in this game can hinder the rest of the group's confidence about coordination in the future rounds.

Session 5 consisted of 12 subjects and only one group. In this experiment, the cost was varied from $c = -0.05$ to $c = 1.05$ and back down to $c = 0$. In the first half of the experiment, the subjects

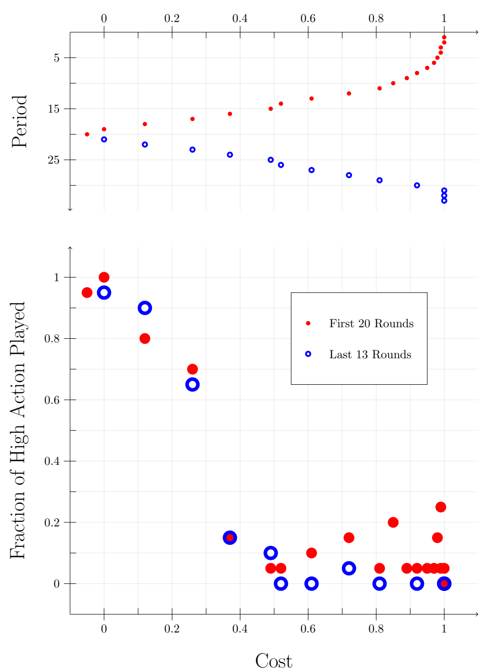
[1] This information was obtained in a questionnaire that was given after the experiment.



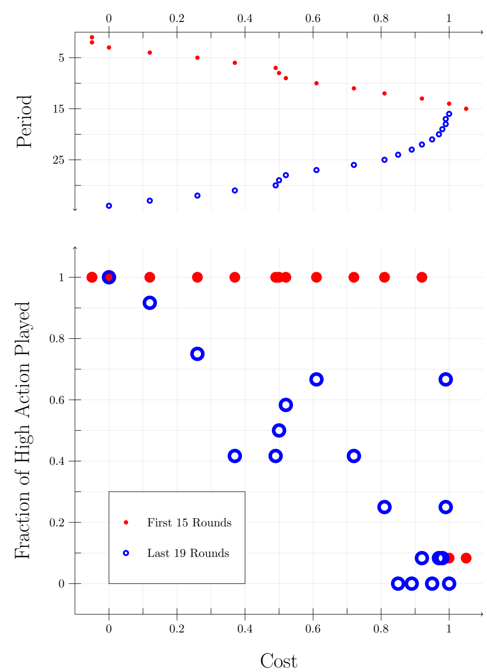
(a) Session #2



(b) Session #3



(a) Session #4



(b) Session #5

Figure 4.9: Experimental Results.

were able to coordinate perfectly on the high-effort action until the $c = 1$ game, at which point all but one subject switched to the low-effort action. Then as the cost started to decrease, the majority of the subjects remained playing the low-effort action, including 4 rounds where all 12 of the subjects played the low-effort action. As the cost started to decrease more, the subjects gradually started switching back to the high-effort action, and finally reached consensus on the high-effort action at the $c = 0$ game. The experimental results for session 5 are displayed in Figure 4.9(d).

This session showed the hysteresis that was hypothesized can happen in coordination games. For example, in the 13^{th} round, at a cost of $c = 0.92$, 12 out of 12 subjects played the high-effort action. However, 10 rounds later, 11 out of 12 subjects played the low-effort action at the same cost. These two rounds had the exact same game, payoffs, and group, and were played within a couple minutes of each other; yet the equilibrium in these two games were completely opposite. The shock occurred when the system reached the $c = 1$ game at which point almost everyone switched from the high-effort action to the low-effort action. Then as the cost parameter decreased, the equilibrium outcome took time to adjust back to the original solution, which is exactly the hysteresis that was hypothesized.

The round-by-round outcomes of session 5 are displayed in Figure 4.10. The solid circles denote the outcome for the first 15 rounds and the empty circles are the outcomes for the last 19 rounds. For the first 13 rounds, the only outcome is the high-effort outcome. In the 14^{th} round the cost reaches $c = 1$, and the low-effort outcome ensues for the next 17 rounds. Finally in the last round when the cost gets back down to $c = 0$, the group outcome changes back to the high-effort outcome.

In terms of maximum likelihood, the data does not match the predicted curves very well. However, this is from a small data set, and therefore may fit better if more data is collected. Whether this data fits the predicted estimates well or not, it still exhibits hysteresis which was the main goal of the experiment. Even if the theory is not able to perfectly explain this type of behavior, it is good enough to give us insight that can help us discover unintuitive behavior like that found in the experiment.

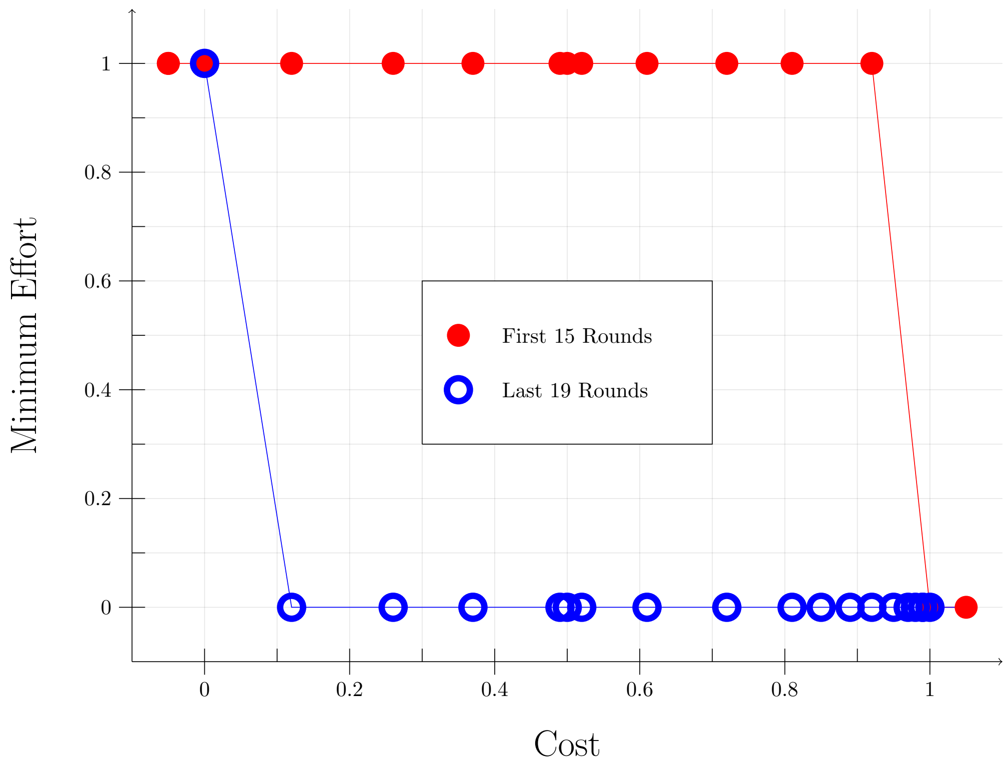


Figure 4.10: Minimum Effort from Session #5.

4.4 Conclusion

This paper started with a fundamental concept from a physical system, and then uncovered the same behavior in an economic system. The similarities in the economic system and the physical system lead to the hypothesis that the economic system would likely exhibit the hysteresis that was present in the physical system. Experiments were used to confirm this hypothesis. The experiment did not match the theory perfectly, but it did confirm that there is hysteresis in the economic system.

It is nice to compare an economic system to a physical system, but what are the substantive implications of this behavior? Consider a multi-period public good model, where people can either contribute or not contribute to the public good in every period. This is a classical coordination game with two equilibria, one where everyone contributes and one where no one contributes. Given certain levels of costs and benefits, it is likely that the socially optimal outcome will be the equilibrium where everyone contributes. However, it is also likely that the actual outcome is the equilibrium where no one contributes, because it is a collective action problem. Given that this system has the same double saddle-node bifurcation that the above coordination game does, varying the cost could induce the group to switch from the lower equilibrium to the higher equilibrium while still maintaining a balanced budget. To do this, a one period cost subsidy would be given to the group, making the cost of the public good less than normal. If this subsidy was large enough, then the group would reach the socially optimal equilibrium because individuals would start to contribute if they received this subsidy. Then in the next period, increase the price above the original price in order to regain the lost money from the subsidy. Since the system exhibits hysteresis, the equilibrium would stay at the high outcome, even when the cost is increased higher than the original value. Then once the subsidy has been gained back, the cost could be set back to the original point, and the group would be at the socially optimal equilibrium, and the government's budget would be balanced.

Examples like the above show the importance of understanding the behavior of bifurcation correspondences in equilibrium models. Even though session 5 showed the desired result, the other treatments were not successful. As seen by the experimental results, it seems that more than two subjects are needed for this to work, especially with Caltech subjects. This is consistent with the finding that this double saddle-node bifurcation is more likely the larger the group size. In the context of the above example, this is good, because larger groups have worse collective action problems.

Another important implication of this result is that in general, the order of experiments is very important. In session 5, the subjects played the exact same game, with the same group twice within 2 minutes of each other. However, the second time, the subjects played completely opposite than they did the first time. This suggests that when running multiple rounds in one treatment for an experiment, the order of the experiment is incredibly important.

After examining the theory, formulating hypotheses, and confirming these experimentally, there are still many extensions. First, it would be helpful to collect more data, and gain a better understanding of these coordination games. It would be useful to run some experiments away from Caltech, where the subject pool is more heterogeneous. Also, the experimental data did not fit the theoretical predictions well. If after further experimentation, the data still does not fit the theory well, then the theory may need modifications to help explain this type of behavior.

Appendices

Appendix A

Computational Testbeds for Coordination Games

A.1 Simulation Parameters

A.1.1 FP Simulations

The initial choice in the fictitious play simulations is chosen from $U[0,1]$. After this players best respond to empirical distribution of the history of play. In the continuous minimum-effort coordination game, the best response to any history of choices will be either 0, 1 or one of the previous choices. To calculate this best response, the players cycle through all possibilities and keep the one that gives the highest payoff.

A.1.2 AL Simulations

In the simulations from Arifovic and Ledyard's (2005) algorithm, each agent keeps a collection of choices. The size of the collection is $J = 50$. The collection is initialized by drawing J numbers from $U[0,1]$. At each round, the agent's action is chosen randomly from the collection based on a probability distribution weighted by the forgone utility[1] of each action. The agents then update their choice set by experimentation and replication of the good strategies. The experimentation parameter is $\rho = 0.03$. After the choice set is updated, the probability distribution is updated for the following round, and then the agent is ready to make his choice for the next round. A further

[1]What the payoff would have been if the agent had chosen this instead of what they chose the last round.

explanation of the AL algorithm is available in their paper.

A.1.3 PR Simulations

The PR simulations are described in detail in the body of the paper. All of the simulations used the following parameters. Agents take $J = 10$ random samples when updating their quasi-best-response and their highest and lowest known payoffs. The boundaries for the confidence parameter are $\gamma \in [10, 10000]$. The parameter relating variance and confidence is set to $\rho = 12$. When updating the confidence, the agents have $\alpha_1 = 1/3$ and $\alpha_2 = 1/6$. This means that the agents' confidence increases when their prediction is within $1/6$ of the actual outcome. When the agents are recognizing patterns, they look for the $j = 3$ most similar plays from the past to make their prediction. Finally, the agents won't start a moving session unless their confidence parameter is at least $\gamma_{MS} = 50$. Once the agents are in a moving session, they are allowed $\bar{f} = 4$ failures before they become discouraged and stop the moving session.

A.2 Experimental Instructions

Procedural Summary

There are five papers folded in half on each of your desk. They are labeled "Table 1", "Table 2", Table 3", "Table 4" and "Table 5". Each of these papers will present a different payoff table. The payoff from your paper labeled "Table 1" will be identical to everyone else's paper labeled "Table 1", the payoff from your "Table 2" will be identical to everyone else's "Table 2" and so on. Please do not open any of these papers until you are instructed to do so.

The experiment consists of 5 blocks and each block will consist of 15 periods. You will be placed in a group with 3 others who are randomly chosen for each block. This means that at the beginning of each block, you will be randomly matched with 3 others individuals to be placed in a group size of 4.

Your payoff from each period will be determined by one of the 5 folded papers. You will be using one payoff table for the entire block. Everyone in your group will be using the exact same payoff table as you. Your group members are randomly chosen (via random number generator) and your payoff tables for each block are also randomly chosen (via random number generator). Your payment will be sum of your entire earnings from all 5 blocks.

Q: How do you know what period you are in? What Payoff table to use? Which block you are in?
A: Refer to Figure 1. Number of periods is denoted in the upper left corner. The first line of the computer's instruction tells you what table your payoff is determined from.

Periods 1-15 is block one. Periods 16-30 is block two. Periods 31-45 is block three. Periods 46-60 is block four. Periods 61-75 is block five.

Note: There will be no sign that tells you that you are starting a new block. So pay attention to the period numbers.

Q: Is it possible to have the same table number from one block to another?
A: Yes, this is because you are randomly assigned a payoff table for each block.

Timeline and Summary

Block 1 begins. You're randomly matched with 3 other people and everyone in your group is randomly assigned a same payoff table to use for this block. You may open the payoff table at this time. The period will begin and all the participants will make their choices privately through the computer. Press "okay" after you have made your decision. After all the participants have inputted their choices, the lowest number chosen from your group as well as your payoff will be displayed. Press continue to move on to the next period. You and your group will do this for 15 periods and block 1 will come to an end.

Block 2 begins. You're randomly matched with 3 other people and everyone in your group is randomly assigned a same payoff table to use for this block. You may open the payoff table at this time. The period will begin and all the participants will make their choices privately through the computer. Press "okay" after you have made your decision. After all the participants have inputted their choices, the lowest number chosen from your group as well as your payoff will be displayed. Press continue to move on to the next period. You and your group will do this for 15 periods and block 2 will come to an end.

This continues until all 5 blocks are played. Your payment will be the sum of your payoff from each period in all 5 blocks.

Instructions - Page 1

Experiment Overview

You are about to participate in an experiment in the economics of decision making. If you listen carefully and make good decisions, you could earn a considerable amount of money that will be paid to you in cash at the end of the experiment.

Please do not talk or communicate with other participants. Feel free to ask questions by raising your hand or signaling to the experimenter.

You will be working with a fictitious currency called Francs. The exchange rate will be specified in the instructions. You will be paid in cash at the end of the experiment.

The experiment consists of a sequence of periods and blocks. There will be total of 5 blocks. For each block, there will be total of 15 periods.

Specific Instructions for Each Period

Exchange rate: ______ Francs = _____ USD.

Your group will consist of you and 3 other individuals (total of 4 people in your group). Your job is to choose one of the following numbers: {1, 2, 3, 4, 5, 6, 7}. The number you choose will remain anonymous. Your individual payoff is determined by your choice and the choice of others in your group. The following is a sample payoff table for illustration purposes only. Your actual payoff table will be using different numbers from this table. The overall ideal will be the same, however.

Table 1: Your payoff in francs

	Lowest choice of number from your group (including you)						
Your choice of number	7	6	5	4	3	2	1
7	25	21	17	13	9	5	1
6	-	23	19	15	11	7	3
5	-	-	21	17	13	9	5
4	-	-	-	19	15	11	7
3	-	-	-	-	17	13	9
2	-	-	-	-	-	15	11
1	-	-	-	-	-	-	13

Examples
- You chose 5 and the lowest choice of number from your group is 5. Then you win 21 francs.
- You chose 4 and the lowest choice of number from your group is 2. Then you win 11 francs.
- You chose 3 and the lowest choice of number from your group is 2. Then you win 13 francs.

Quiz
You chose 2 and the lowest choice of number from all the participants is 1. Then you win _____ francs.

Any questions?

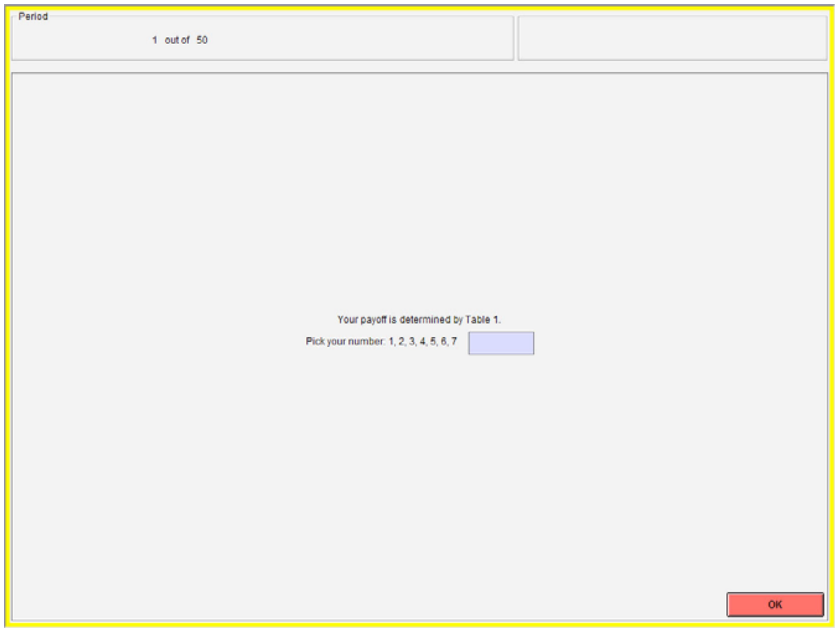


Figure 1: Sample Screenshot

A.3 Tables

	c = 50	c = 500	c = 900	c = 950	c = 990
Choice	4.8485	4.5000	1.2864	1.2606	1.1242
SE	0.0932	0.0975	0.0363	0.0391	0.0244

Table A.1: Average Choice for Different Cost Parameters (Data for Figure 2.10).

	$\mu_{50} > \mu_{500}$	$\mu_{500} > \mu_{900}$	$\mu_{900} > \mu_{950}$	$\mu_{950} > \mu_{990}$
p-value	0.0126	0	0.1677	0.0023
t-value	2.2428	21.2006	−0.9642	2.8422

Table A.2: Average Choice Comparison (Data for Figure 2.10).

	c = 50	c = 500	c = 900	c = 950	c = 990
μ	9088	6527	4968	4846	5136
SE_μ	118.05	103.36	115.42	124.51	106.56

Table A.3: Average Payoffs for Different Cost Parameters After 4 Rounds (Data for Figure 2.14).

	$\mu_{50} > \mu_{500}$	$\mu_{500} > \mu_{900}$	$\mu_{900} > \mu_{950}$	$\mu_{950} < \mu_{990}$
p-value	0	0	0.2366	0.0393
t-value	16.3234	10.05	0.7178	1.7632

Table A.4: Average Payoffs Comparison After 4 Rounds (Data for Figure 2.14).

	c = 50	c = 500	c = 900	c = 950	c = 990
μ	9791	7489	5652	5560	5650
SE_μ	76.23	53.75	38.61	43.66	35.02

Table A.5: Average Payoffs for Different Cost Parameters After 15 Rounds (Data for Figure 2.15).

	$\mu_{50} > \mu_{500}$	$\mu_{500} > \mu_{900}$	$\mu_{900} > \mu_{950}$	$\mu_{950} < \mu_{990}$
p-value	0	0	0.0557	0.0536
t-value	24.6797	27.75	1.5928	1.6114

Table A.6: Average Payoffs Comparison After 15 Rounds (Data for Figure 2.15).

Appendix B

Bounded Rationality in Repeated Games

B.1 Structure of Automata

The set of finite automata contains many automata which are redundant. It simplifies the analysis to eliminate some of these redundant automata, allowing me to focus on a smaller set of automata. Much of the notation from this section is from Kohavi (1978).

B.1.1 Payoff Equivalent Automata

Definition B.1.1 (Payoff Equivalent Automata)**.** *Automata M_1 and M_2 are said to be payoff equivalent over set $\mathcal{M}$ if and only if,*

$$U_i\left(M_1, A, \varepsilon\right) = U_i\left(M_2, A, \varepsilon\right) \text{ for all } A \in \mathcal{M}, \text{ and all } \varepsilon \in (0, .5].$$

Two automata are considered payoff equivalent over a set $\mathcal{M}$ if they yield the same payoff when matched against any automaton from $\mathcal{M}$. For any set of payoff equivalent automata $\mathcal{M}^{PE}$, I only need to consider one automaton $M_1 \in \mathcal{M}^{PE}$ when calculating equilibria. When M_1 is not part of an equilibrium, none of the automata in $\mathcal{M}^{PE}$ are part of an equilibrium. When M_1 forms an equilibrium with M_2, then any automaton from $\mathcal{M}^{PE}$ forms an equilibrium with M_2. When computing equilibria in my model, I can without loss of generality search over a smaller set of automata where any set of payoff equivalent automata is represented by a single automaton.

B.1.2 Reduced Automata

Next, I introduce the concept of a reduced automaton. Any non-reduced automaton is payoff equivalent to some reduced automaton. Therefore, I am able to only focus on the set of reduced automata without loss of generality.

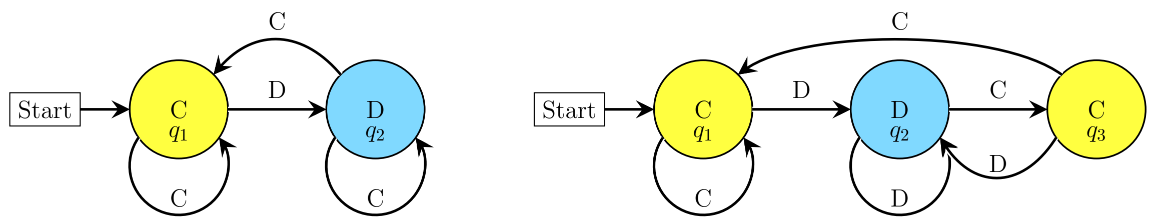
Definition B.1.2 (Equivalent States)**.** *States s_i and s_j are said to be equivalent if and only if, for every possible input sequence, the same output sequence is produced, regardless of whether s_i or s_j is the initial state.*

Definition B.1.3 (Equivalent Automata)**.** *Two automata, M_1 and M_2, are said to be equivalent if and only if, for every state in M_1, there is a corresponding equivalent state in M_2, and vice versa.*

If two automata are equal, then they must be equivalent. However, if two automata are equivalent they need not be equal. Each of the automata in Figure B.1 represent the tit-for-tat strategy. Figure B.1(a) is a two-state automaton which represents tit-for-tat, while Figure B.1(b) is three-state automaton which represents tit-for-tat. Both q_1 and q_3 from B.1(b) are equivalent to q_1 from B.1(a), and state q_2 in B.1(b) is equivalent to q_2 in B.1(a), so these automata are equivalent but not equal.

Definition B.1.4 (Reduced Automaton)**.** *An automaton M is reduced if and only if it contains no equivalent states.*

Every non-reduced automaton has a corresponding reduced automaton, where equivalent states are combined into a single state. The non-reduced automata and the corresponding reduced automata are payoff equivalent over the set of finite automata, because they produce the same output for all sequences of input. I am therefore able to restrict the set of automata from all finite automata to reduced automata without loss of generality.



(a) 2-state Tit-for-Tat (b) 3-state Tit-for-Tat

Figure B.1: Example of equivalent but not equal automata.

B.1.3 Strongly Connected Automata

Next, I introduce the notion of a strongly connected component, an absorbing region of the automaton.

Definition B.1.5 (Reachable State)**.** *Given automaton $M = (Q, q^0, f, \tau)$, state $q_m \in Q$ is reachable from $q_1 \in Q$ if there exists some sequence of signals, $\boldsymbol{r} = \{r_1, \ldots, r_m\}$ such that,*

$$\tau(q_k, r_k) = q_{k+1} \text{ for all } 1 \leq k \leq m-1 ,$$

where states $q_2, \ldots, q_{m-1}$ are defined recursively.

Definition B.1.6 (Strongly Connected Subset)**.** *Given automaton $M = (Q, q^0, f, \tau)$, a subset of states $Q^{SCS} \subseteq Q$ is said to be strongly connected if for every pair of states $q_i, q_j \in Q^{SCS}$, q_i is reachable from q_j.*

Definition B.1.7 (Strongly Connected Component)**.** *Given automaton $M = (Q, q^0, f, \tau)$, a subset of states $Q^{SCC} \subseteq Q$ is said to be strongly connected component (SCC) if Q^{SCC} is strongly connected and there is no state $q \in Q \backslash Q^{SCC}$ such that $Q^{SCC} \cup q$ is strongly connected.*

A strongly connected component is a region of the automaton that cannot be left once it has been reached regardless of the future signal sequence. All states in a strongly connected component are reachable from all other states in the SCC. Therefore, once the automaton enters one of these SCCs, all other states of the automaton become irrelevant.

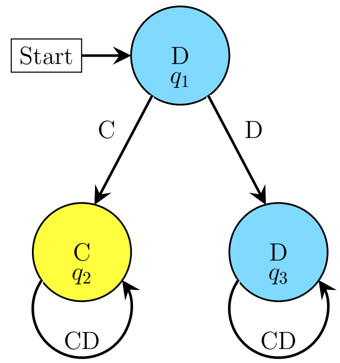


Figure B.2: Non-strongly-connected automaton.

Definition B.1.8 (Strongly Connected Automaton)**.** *Automaton* $M = (Q, q^0, f, \tau)$ *is strongly connected if* Q *is a strongly connected component.*

An example of an automaton that is not strongly connected is displayed in Figure B.2. This automaton has three states, $Q = \{q_1, q_2, q_3\}$. It is clear by definition that this automaton has two strongly connected components; $Q_1^{SCC} = \{q_2\}$ and $Q_2^{SCC} = \{q_3\}$, and therefore is not a strongly connected automaton. The automaton starts in state q_1. If it receives a C signal in the first round, then it enters q_2 and always plays C. If it receives a D signal in the first round, then it enters q_3 and always plays D. So with certain probability this automata always plays C, otherwise it always plays D.

Every automaton has at least one strongly connected component. When signal are imperfect, the automaton reaches a SCC with probability one, and remains in that SCC for the remainder of the supergame. Since I am focusing on the long run behavior of the automata, I restrict the set of automata to only strongly connected automata. It is important to note that if player 1 plays a strongly connected automaton M_1, then player 2 is at least weakly better playing a strongly connected automaton as well.

Lemma B.1.9. *For* $M_1 \in \mathcal{M}^{SCC}$ *and* $M_2 \in \mathcal{M} \backslash \mathcal{M}^{SCC}$ *and any* $\varepsilon \in (0, .5]$*, there exists* $M_2' \in \mathcal{M}^{SCC}$ *such that* $U_2(M_1, M_2', \varepsilon) \geq U_2(M_1, M_2, \varepsilon)$*.*

Therefore, any equilibrium over the set of strongly connected automata is also an equilibrium over the set of all finite automata. However, there may be equilibria that contain one or more automata which are not strongly connected.

The idea for this proof is as follows. Suppose player 1 plays a strongly connected automaton. If player 2 plays an automaton with more than one strongly connected component, then depending on the starting state, the system may enter any one of the strongly connected components with positive probability. If different strongly connected components yield different payoffs, player 2 is better playing the automaton with only the strongly connected component with the highest payoff.

To summarize, for the N-state analysis, I restrict the set of finite automata to those which are finite, strongly connected, and reduced. This set is denoted by $\mathcal{M}_i^R$. All equilibria over this set are also equilibria over the set of all finite automata. However, there may be additional equilibrium consisting of one or more non-strongly connected automata.

Proof of Lemma B.1.9

Since $M_2 = \left(Q_2, q_2^0, f_i, \tau_i\right)$ is not a strongly connected automaton, then the states can be divided up into strongly connected components and transient classes. Let $Q_1^{SCC}, \ldots, Q_n^{SCC} \subset Q_2$ be the strongly connected components of automaton M_2.

First, consider the trivial case that automaton starts in a strongly connected component, $q^0 \in Q_k^{SCC}$. Then let automaton M_2' have the states Q_k^{SCC} and the corresponding output function, transition function, and starting points from M_2. Since Q_k^{SCC} is strongly connected component this automaton is well defined. It is clear by the definition of strongly connected components that M_2 and M_2' yield the same payoff against M_1.

Next, consider the situation where M_2 does not start in a strongly connected component, $q_2^0 \notin Q_k^{SCC}$ for any $k = 1, \ldots, n$. Given the starting point x^0 corresponding to q_1^0 and q_2^0, the system $X\left(M_1, M_2, \varepsilon\right)$ has a unique stationary distribution $\pi\left(M_1, M_2, \varepsilon\right)\left(x^0\right)$. This stationary distribution is the convex combination of stationary distributions,

$$\pi\left(M_1, M_2, \varepsilon\right)\left(x^0\right) = \sum_{k=1}^{n} \beta_k \pi_k,$$

where β_k is the probability that starting at x^0 the system gets absorbed to Q_k^{SCC}, and π_k is the stationary distribution of the system when M_2 starts in Q_k^{SCC}. The payoff is therefore written as

$$U_2(M_1, M_2, \varepsilon) = \sum_{k=1}^{n} \beta_k U_2\left(M_1, M_k^{SCC}, \varepsilon\right),$$

where M_k^{SCC} is the automaton composed of the states Q_k^{SCC}. Let M_2' be the automaton M_k^{SCC} which yields the highest payoff against M_1 and has positive probability of being reached, $\beta_k > 0$. Then this automaton yields at least weakly higher payoffs against M_1 than M_2. ■

B.2 Proofs

I present the finite-state results first, as some of these are used in the two-state results.

B.2.1 Finite-State Results

Lemma B.2.1. *Given $M_1 \in \mathcal{M}_1^R$ and $M_2 \in \mathcal{M}_2^R$ and regular signal functions $r_i(s_i, s_j, \varepsilon)$, then the Markov chain $X(M_1, M_2, \varepsilon)$ is irreducible for all $\varepsilon > 0$.*

Proof of Lemma B.2.1

The Markov chain starts in state x^0, the state corresponding to the situation where both automata are in their initial state, $q_1\left(x^0\right) = q_1^0$ and $q_2\left(x^0\right) = q_2^0$. By definition, the Markov chain has one state for each automata-state pair that is reachable from x^0 with positive probability. Therefore,

$$\left[X(M_1, M_2, \varepsilon)\left(x^0, x\right)\right]^N > 0 \text{ for all } x \in X(M_1, M_2, \varepsilon), \text{ all } \varepsilon > 0, \text{ some } N \geq 0.$$

So every state is reachable from x^0. Next, I show that x^0 is reachable from every state $x \in X(M_1, M_2, \varepsilon)$. By definition, an automaton $M_i = \left(Q_i, q_i^0, f_i, \tau_i\right)$ is strongly connected if Q_i is a strongly connected component. This means that every state in Q_i is reachable from every other state. Therefore, there is some sequence of actions, $\boldsymbol{s}_i(q_1, q_2)$, which takes M_i from state $q_1 \in Q_i$ to state $q_2 \in Q_i$. By the second condition of regular signal function, for all $\varepsilon > 0$, the probability that player i sees sequence of signals $\boldsymbol{s}_i(q_1, q_2)$ is greater than 0.

I want to show that it is possible to get from any state $x \in X(M_1, M_2, \varepsilon)$ to state x^0. Let $q_i(x)$ be the state of M_i when $X(M_1, M_2, \varepsilon)$ is in state x. Then there exists sequences of actions $\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)$, $\boldsymbol{s}_1\left(q_1(x), q_1(x^0)\right)$, $\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)$, and $\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)$. Since x is reachable from x^0, then there exists sequences $\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)$ and $\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)$ of equal length, $\left|\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)\right| = \left|\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)\right|$. The length of the other sequences may not be equal.

If player 2 plays sequences $\boldsymbol{s}_1\left(q_1(x), q_1(x^0)\right)$ and $\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)$ repeatedly $\left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| + \left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| - 1$ times, and then $\boldsymbol{s}_1\left(q_1(x), q_1(x^0)\right)$ is played one more time, then M_1 goes from $q_1(x)$ to $q_1(x^0)$ in

$$\left(\left|\boldsymbol{s}_1\left(q_1(x), q_1(x^0)\right)\right| + \left|\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)\right|\right)\left(\left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| + \left|\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)\right|\right) - \left|\boldsymbol{s}_1\left(q_1(x^0) - q_1(x)\right)\right|$$

moves. Similarly, if player 1 plays sequences $\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)$ and $\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)$ repeatedly $\left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| + \left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| - 1$ times, and then $\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)$ is played one more time, then M_2 goes from $q_2(x)$ to $q_2(x^0)$ in

$$\left(\left|\boldsymbol{s}_1\left(q_1(x), q_1(x^0)\right)\right| + \left|\boldsymbol{s}_1\left(q_1(x^0), q_1(x)\right)\right|\right)\left(\left|\boldsymbol{s}_2\left(q_2(x), q_2(x^0)\right)\right| + \left|\boldsymbol{s}_2\left(q_2(x^0), q_2(x)\right)\right|\right) - \left|\boldsymbol{s}_2\left(q_2(x^0) - q_2(x)\right)\right|$$

moves. The length of these sequences are the same. So each automaton goes from $q_i(x)$ to $q_i(x^0)$, meaning the system goes from x to x^0 with positive probability. So the Markov chain is irreducible.■

Lemma 3.5.3. *Suppose players play automata M_1 and M_2. The average payoff for the infinitely repeated game is equal to*

$$U_i(M_1, M_2, \varepsilon) = \sum_{x_k \in X(M_1, M_2, \varepsilon)} \pi(M_1, M_2, \varepsilon)(x_k)\, u_i(x_k),$$

where $\pi(M_1, M_2, \varepsilon)(x_k)$ is the term of the stationary distribution corresponding to state x_k, and $u_i(x_k)$ is the payoff for player i state x_k.

Proof of Lemma 3.5.3

By Lemma B.2.1, $X(M_1, M_2, \varepsilon)$ is irreducible and hence has a unique stationary distribution $\pi(M_1, M_2, \varepsilon)$ for all $\varepsilon > 0$. Let $H(x_i, T) = \frac{1}{T}\sum_{t=0}^{T} I\{x^t = x_i\}$ be the number of times that $X(M_1, M_2, \varepsilon)$ has visited state x_i in T rounds. By the law of large numbers for irreducible Markov chains (Theorem 11.12 p.439 (Grinstead and Snell 1997)), for all starting states,

$$\lim_{T\to\infty} H(x_k, T) = \pi(M_1, M_2, \varepsilon)(x_k),$$

where $\pi(M_1, M_2, \varepsilon)(x_k)$ is the term of $\pi(M_1, M_2, \varepsilon)$ corresponding to state x_k. The payoff for the first T rounds can be rewritten as,

$$U_i^T(M_1, M_2, \varepsilon) = \sum_{x_k \in X(M_1, M_2, \varepsilon)} H(x_k, T)\, u_i(x_k).$$

Therefore,

$$U_i(M_1, M_2, \varepsilon) = \lim_{T\to\infty} \sum_{x_k \in X(M_1, M_2, \varepsilon)} H(x_k, T)\, u_i(x_k) = \sum_{x_k \in X(M_1, M_2, \varepsilon)} \pi(M_1, M_2, \varepsilon)(x_k)\, u_i(x_k).$$

■

The infinite set of all possible absorbing classes of automaton M is denoted by $AC(M)$. The set of payoff-maximal absorbing states for player i is,

$$AC_i^*(M) = \{a | U_i(a) \geq U_i(b) \text{ for all } b \in AC(M)\}.$$

Lemma B.2.2. *If $a = \{\boldsymbol{q}, \boldsymbol{s}\} \in AC_i^*(M)$ and $q_j, q_k \in \boldsymbol{q}$ such that $q_j = q_k$, then there exists $a' \in AC_i^*(M)$ such that $|a'| < |a|$.*

Proof of Lemma B.2.2

Consider absorbing classes $a = \{\boldsymbol{q}, \boldsymbol{s}\} \in AC_i^*(M)$ with $q_j, q_k \in \boldsymbol{q}$ such that $q_j = q_k$. Then consider the two absorbing classes:

$$a_1 = \left(\{q_1, \ldots, q_{j-1}, q_j, q_{k+1}, q_n\}, \{s_1, \ldots, s_{j-1}, s_k, s_{k+1}, \ldots, s_n\}\right)$$

and

$$a_2 = \left(\{q_{j+1}, \ldots, q_k\}, \{s_{j+1}, \ldots, s_{k-1}, s_j\}\right).$$

Both of these satisfy the conditions for an absorbing class, because,

$$\tau\left(q_j, s_k\right) = \tau\left(q_k, s_k\right) = q_{k+1}$$

and

$$\tau\left(q_k, s_j\right) = \tau\left(q_j, s_j\right) = q_{j+1}.$$

The payoff of absorbing class a is,

$$\begin{aligned}
U_i^{AC}(a) &= \frac{1}{n}\sum_{l=1}^{n} u_i\left(f\left(q_l\right), s_l\right) \\
&= \frac{1}{n}\left[(n-k+j)\left(\frac{1}{n-k+j}\sum_{l=1,k+1}^{j,n} u_i\left(f\left(q_l\right), s_l\right)\right) + (k-j)\left(\frac{1}{k-j}\sum_{l=j+1}^{k} u_i\left(f\left(q_l\right), s_l\right)\right)\right] \\
&= \left(\frac{n-k+j}{n}\right)U_i^{AC}\left(a_1\right) + \left(\frac{k-j}{n}\right)U_i^{AC}\left(a_2\right).
\end{aligned}$$

Since $a \in AC_i^*(M)$, it must be that $U_i^{AC}(a) = U_i^{AC}\left(a_1\right) = U_i^{AC}\left(a_2\right)$, or else either a_1 or a_2 would have higher payoff than a. Since $0 < j < k < n$, $|a_1| = n-k+j < n = |a|$ and $|a_2| = k-j < n = |a|$. So for all payoff-maximal absorbing classes with multiple visits to one state, there exists a smaller payoff-maximal absorbing class. ∎

The set of absorbing classes which contain all unique states for player i is denoted by,

$$AC_i^U(M) = \{a | q_i \neq q_j \text{ for all } q_i, q_j \in a\}.$$

Lemma B.2.3. *At least one of the unique state absorbing classes is payoff-maximal, i.e.*

$$AC_i^*(M) \cap AC_i^U(M) \neq \emptyset.$$

Proof of Lemma B.2.3

Select any absorbing class $a \in AC_i^*(M)$. By Lemma B.2.2, for each payoff-maximal absorbing class that visits state q_j more than once, there exists another payoff-maximal absorbing class that visits q_i strictly less. This process can be repeated until the new absorbing class visits state q_j only once. This can be done for each state in a. Then end result is a payoff-maximal absorbing class in the set of unique state absorbing classes. ■

Lemma B.2.3 suggests that there is a payoff-maximal absorbing class with weakly fewer states than automaton M, which means that there is a finite optimal absorbing class. Player i's payoff-maximal absorbing class is denoted by $a_i^*(M)$.

Lemma B.2.4. $U_i(M_1, M_2, \varepsilon) \leq U_i^{AC}(a_i^*(M_{-i}))$ *for all* $\varepsilon \in [0, .5]$ *and all* $M_i \in \mathcal{M}$.

Proof

Suppose by means of contradiction that for some $\varepsilon \in [0, .5]$, $U_2(M_1, M_2, \varepsilon) > U_2^{AC}(a_2^*(M_1))$. The Markov chain $X(M_1, M_2, \varepsilon)$ yields a sequence of automaton-state pairs $x^0, x^1, x^2, \ldots$. By definition,

$$U_2(M_1, M_2, \varepsilon) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} u_2\left(x^t\right),$$

where $u_2(x^t)$ is the payoff to player 2 for the automaton-state profile x^t. For every finite integer K, there must be some sequence of length K of automaton-state pairs $y^1, \ldots, y^K$ such that,

$$\frac{1}{K}\sum_{k=1}^{K} u_2\left(y^k\right) \geq U_2\left(M_1, M_2, \varepsilon\right). \tag{B.1}$$

Let $|Q_1|$ be the number of states in M_1, and let $\underline{u}_2$ be the lowest possible stage-game payoff for player 2. Set $\bar{K}$ to be a sufficiently high integer such that

$$U_2\left(M_1, M_2, \varepsilon\right) - U_2\left(a_2^*\left(M_1\right)\right) > \frac{j\left(U_2\left(M_1, M_2, \varepsilon\right) - \underline{u}_2\right)}{\bar{K} + j} \tag{B.2}$$

holds for all $j = 1, \ldots, |Q_1|$. From (B.2), we get,

$$\begin{aligned} U_2\left(a_2^*\left(M_1\right)\right) &< U_2\left(M_1, M_2, \varepsilon\right) - \frac{j\left(U_2\left(M_1, M_2, \varepsilon\right) - \underline{u}_2\right)}{\bar{K} + j} \\ &= \frac{\bar{K}U_2\left(M_1, M_2, \varepsilon\right) - j\underline{u}_2}{\bar{K} + j}. \end{aligned} \tag{B.3}$$

Fix sequence of automaton-state pairs $y^1, \ldots, y^{\bar{K}+1}$ such that,

$$\frac{1}{\bar{K}}\sum_{k=1}^{\bar{K}} u_2\left(y^k\right) \geq U_2\left(M_1, M_2, \varepsilon\right).$$

Let $q_i^1, q_i^2, \ldots, q_i^{\bar{K}+1}$ be the sequence of states for automaton i from the sequence of automaton-state pairs $y^1, \ldots, y^{\bar{K}+1}$. Automaton M_1 starts in state q_i^1 and ends in state $q_i^{\bar{K}+1}$. Because M_1 is strongly connected, there exists some sequence of actions $s_2^1, s_2^2, \ldots, s_2^j \in S_2$, that moves automaton M_1 from state $q_i^{\bar{K}+1}$ to state q_i^1. Let $p_1^1 = q_1^{\bar{K}+1}$ and $p_1^l = \tau_2\left(p_1^{l-1}, s_2^{l-1}\right)$ for $l = 2, \ldots j$. By construction it must be that, $\tau_2\left(p_1^j, s_2^j\right) = q_1^1$. Since M_1 has $|Q_1|$ states, then $\left|\left\{s_2^1, s_2^2, \ldots, s_2^j\right\}\right| \leq |Q_1|$. Define the absorbing class

$$a_2'\left(M_1\right) = \left(\left\{q_1^1, q_1^1, \ldots, q_1^{\bar{K}}, p_1^1, p_1^2, \ldots, p_1^j\right\}, \left\{f_2\left(q_2^1\right), f_2\left(q_2^2\right), \ldots, f_2\left(q_2^{\bar{K}}\right), s_2^1, s_2^2, \ldots, s_2^j\right\}\right).$$

This is a well defined absorbing class. Note that $u_2\left(f_1\left(p_1^k\right), s_1^k\right) \geq \underline{u}_2$ for $k = 1, \ldots, j$. Therefore,

$$\begin{aligned}
U_2^{AC}\left(a_2'\left(M_1\right)\right) &= \frac{1}{\bar{K}+j}\left[\sum_{k=1}^{\bar{K}} u_2\left(y^k\right) + \sum_{k=1}^{j} u_2\left(f_1\left(p_1^k\right), s_2^k\right)\right] \\
&\geq \frac{1}{\bar{K}+j}\left[\sum_{k=1}^{\bar{K}} u_2\left(y^k\right) + j\underline{u}_2\right] && \text{(By minimality of } \underline{u}_2.\text{)} \\
&\geq \frac{\left[\bar{K}U_2\left(M_1, M_2, r_1(\varepsilon), r_2(\varepsilon)\right) + j\underline{u}_2\right]}{\bar{K}+j} && \text{(From (B.1))} \\
&> U_2\left(a_2^*\left(M_1\right)\right) && \text{(From (B.3))}
\end{aligned}$$

This contradicts the maximality of $a_2^*\left(M_1\right)$. ■

Lemma B.2.5. *Given regular signal functions r_i, if $X\left(M_1, M_2, \varepsilon\right)$ has communicating classes $A_1, \ldots, A_m$, then,*

$$\lim_{\varepsilon \to 0} U_i\left(M_1, M_2, \varepsilon\right) = \sum_{A_k | \gamma(A_k) = \gamma^*} \beta(A_k) U_i^{CC}\left(A_k\right),$$

with $\sum_{A_k|\gamma(A_k)=\gamma^*} \beta(A_k) = 1$ and $\beta\left(A_k\right) > 0$ for all A_k such that $\gamma\left(A_k\right) = \gamma^*$

Proof

By Lemma B.2.1, the Markov chain $X\left(M_1, M_2, \varepsilon\right)$ is irreducible and has a unique stationary distribution $\pi\left(M_1, M_2, \varepsilon\right)$. Let $\pi\left(M_1, M_2, \varepsilon\right)(x)$ denote the term of the stationary distribution corresponding to state $x \in X\left(M_1, M_2, \varepsilon\right)$. By Theorem B.2.13, if a communicating class doesn't minimize stochastic potential, $\gamma\left(A\right) > \gamma^*$, then,

$$\lim_{\varepsilon \to 0} \pi\left(M_1, M_2, \varepsilon\right)(x) = 0 \text{ for all } x \in A. \tag{B.4}$$

If a communicating class A does minimize stochastic potential, $\gamma(A) = \gamma^*$, then,

$$\lim_{\varepsilon \to 0} \pi\left(M_1, M_2, \varepsilon\right)(y) > 0 \text{ for all } y \in A. \tag{B.5}$$

For each communicating class A_k, there exists some constant, $\alpha(A_k)$ such that,

$$\lim_{\varepsilon\to 0}\sum_{x\in A_k}\pi(M_1,M_2,\varepsilon)(x)u_i(x)=\alpha(A_k)\sum_{x\in A_k}u_i(x), \tag{B.6}$$

Then,

$$\begin{aligned}
\lim_{\varepsilon\to 0}U_i(M_1,M_2,\varepsilon) &= \lim_{\varepsilon\to 0}\sum_{x\in X(M_1,M_2,\varepsilon)}\pi(M_1,M_2,\varepsilon)(x)u_i(x) && \text{(by Lemma 3.5.3)}\\
&= \lim_{\varepsilon\to 0}\sum_{A_k|\gamma(A_k)=\gamma^*}\sum_{x\in A_k}\pi(M_1,M_2,\varepsilon)(x)u_i(x) && \text{(by (B.4))}\\
&= \sum_{A_k|\gamma(A_k)=\gamma^*}\alpha(A_k)\sum_{x\in A_k}u_i(x) && \text{(by (B.6))}\\
&= \sum_{A_k|\gamma(A_k)=\gamma^*}\alpha(A_k)\,|A_k|\,U_i^{CC}(A_k). && \text{(by def. of } U_i^{CC}\text{)}
\end{aligned}$$

Set $\beta(A_k)=\alpha(A_k)|A_k|$, then $\sum_{A_k|\gamma(A_k)=\gamma^*}\beta(A_k)=\sum_{x\in X(M_1,M_2,\varepsilon)}\pi(M_1,M_2,\varepsilon)(x)=1$ and $\beta(A_k)>0$ for all A_k such that $\gamma(A_k)=\gamma^*$ by (B.5). ■

Definition B.2.6 (Homing Sequence)**.** *Given automaton $M=(Q,q^0,f,\tau)$, the action sequence $h\in S^n$ is a homing sequence if and only if,*

$$\forall q_1,q_2\in Q \text{ and } q_1\langle h\rangle=q_2\langle h\rangle\Rightarrow q_1h=q_2h,$$

where $q\langle h\rangle\in f\left(S^{n+1}\right)$ is the output of M starting at state q when the sequence h is played, and qh is the end state of M when h is played.

This means that when h is played, the output of M allows us to determine the current state of M.

Theorem B.2.7 (Kohavi (1978))**.** *A preset homing sequence, whose length is at most $(n-1)^2$, exists for every reduced, strongly connected n-state machine M.*

Lemma 3.5.7. *Given automaton $M_1 \in \mathcal{M}^R$ with n states, and any absorbing class $a(M_1)$, there exists automaton M_2 such that for all communicating classes, A_k, of the system $X(M_1, M_2, \varepsilon)$,*

$$U_2^{CC}(A_k) = U_2^{AC}(a(M_1)).$$

Proof of Lemma 3.5.7

I construct automaton $M_2 = (Q_2, q_2^0, f_2, \tau_2)$ which yields the desired properties. Consider automaton M_1 with n states and absorbing class

$$a = (\{q_1, \ldots, q_m\}, \{s_1, \ldots, s_m\})$$

with $m \leq n$ states.

The automaton will be made up of three main parts. The first part is the absorbing class. This section of the automaton will keep the system in the desired absorbing state when reached. Then second part of the automaton is the homing region. In this region, the automaton plays the homing sequence. Based on the response from M_1, the current state of M_1 is determined. The goal of the homing region is to determine the current state of automaton M_1 after an error has been made. Once the state of automaton M_1 is known, it will be possible to move it back into the absorbing class. The third part allows the two automata to resynchronize, transitioning from the homing region back to the absorbing class.

Start constructing M_2 by creating states and transitions such that the absorbing class is maintained. That is for each state q_j in the absorbing class a of M_1, create corresponding state p_j in automaton M_2 that satisfies,

$$f_2(p_j) = s_j \text{ and } \tau(p_j, f_1(q_j)) = \begin{cases} p_{j+1} & j < m \\ p_1 & j = m \end{cases}.$$

Also, let all incorrect plays in the absorbing class states lead to state p_{m+1},

$$\tau\left(p_j, s \neq f_1\left(q_j\right)\right) = p_{m+1}.$$

The second region of the automaton is the homing region. By Theorem B.2.7, there exists a homing sequence for automaton M_1, call this $h\left(M_1\right) = \{h_1, \ldots, h_l\}$. There is a set of sequences imposed by this homing sequence when started at different states,

$$S(h) = \{q\langle h\rangle | q \in Q\} = \left\{\begin{array}{c} \left(s_1^1, \ldots, s_l^1\right) \\ \left(s_1^2, \ldots, s_l^2\right) \\ \vdots \\ \left(s_1^k, \ldots, s_l^k\right) \end{array}\right\}.$$

There is also a set of states that this homing sequence will lead to,

$$Q(h) = \{q | q'h = q \text{ for some } q' \in Q\}.$$

Let $S\left(h, j\right)$ for $0 < j \leq l$ be the first j terms of these sequences,

$$S\left(h, j\right) = \left\{\begin{array}{c} \left(s_1^1, \ldots, s_j^1\right) \\ \left(s_1^2, \ldots, s_j^2\right) \\ \vdots \\ \left(s_1^k, \ldots, s_j^k\right) \end{array}\right\} = \left\{\begin{array}{c} \boldsymbol{s}^1(j) \\ \boldsymbol{s}^2(j) \\ \vdots \\ \boldsymbol{s}^k(j) \end{array}\right\}.$$

Let $S^U(h,j)$ be the set of unique sequences of length j imposed by h,

$$S^U(h,j) = \left\{ \begin{array}{c} \left(s_1^1, \ldots, s_j^1\right) \\ \\ \left(s_1^2, \ldots, s_j^2\right) \\ \vdots \\ \\ \left(s_1^u, \ldots, s_j^u\right) \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{s}^1(j) \\ \\ \boldsymbol{s}^2(j) \\ \vdots \\ \\ \boldsymbol{s}^u(j) \end{array} \right\},$$

where $\left|S^U(h,j)\right| = u(j)$ is the number of unique sequences in $S(h,j)$.

The homing region consists of $l+1$ classes, $P_{m+1}, \ldots, P_{m+l+1}$. Class P_{m+1+j} consists of $u(j)$ states, $p_{m+1+j}\left(\boldsymbol{s}^1(j)\right), \ldots, p_{m+1+j}\left(\boldsymbol{s}^{u(j)}(j)\right)$, one corresponding to each sequence in $S^U(h,j)$. Define $S^U(h,0) = \{\emptyset\}$ and $u(0) = 1$. Automaton M_2 will play the same action in all states of a given class,

$$f_2(p) = h_i \text{ for all } p \in P_{m+i} \text{ for } i \in \{1, \ldots, l\}.$$

This choice will correspond to the matching term in the homing sequence.

The transition function for $0 < i \leq l$ is defined as follows.

$$\tau_2\left(p_{m+i}(\boldsymbol{s}), C\right) = \begin{cases} p_{m+i+1}\left(\{\boldsymbol{s}, C\}\right) & \text{if } \{\boldsymbol{s}, C\} \in S^U(h,i) \\ \\ p_{m+i+1}\left(\{\boldsymbol{s}, D\}\right) & \text{if } \{\boldsymbol{s}, C\} \notin S^U(h,i) \end{cases}$$

$$\tau_2\left(p_{m+i}(\boldsymbol{s}), D\right) = \begin{cases} p_{m+i+1}\left(\{\boldsymbol{s}, D\}\right) & \text{if } \{\boldsymbol{s}, D\} \in S^U(h,i) \\ \\ p_{m+i+1}\left(\{\boldsymbol{s}, C\}\right) & \text{if } \{\boldsymbol{s}, D\} \notin S^U(h,i) \end{cases}.$$

Finally, the last region of M_2 will resynchronize play, and get the system back to the absorbing class a. There will be k states in class P_{m+l+1}. By definition of the homing sequence, for each state $p_{m+l+1}(q) \in P_{m+l+1}$ there is a corresponding state $q \in M_1$ such that when M_2 is in state $p_{m+l+1}(q)$, M_1 is in state q. Define the resynchronizing sequence $\boldsymbol{t}(q) = \left\{t_1(q), t_2(q), \ldots, t_{r(q)}(q)\right\}$ to be the sequence of plays necessary to get from state q to state q_1 where $r(q) = |\boldsymbol{t}(q)|$. This sequence exists for each state because M_1 is strongly connected. Then for each state $p^1(q) = p_{m+l+1}(q) \in P_{m+l+1}$, for $0 < i < r(q)$.

$$\tau_2\left(p^i(q), \text{C or D}\right) = p^{i+1}(q)$$

and

$$\tau_2\left(p^{r(q)}(q), \text{C or D}\right) = p_1.$$

The output function for $0 < i \leq r(q)$ is,

$$f_2\left(p^i(q)\right) = t_i(q).$$

So the system will always end up in state (q_1, p_1) regardless of the starting position. Once the system is in (q_1, p_1), it has entered the communicating class, and will not leave without errors. ∎

Definition B.2.8 (Regular Perturbation)**.** *Given Markov chain X, a perturbation X^ε is called a regular perturbation if the following three conditions hold,*

1. *X^ε is irreducible for all $\varepsilon \in (0, .5]$.*
2. $\lim_{\varepsilon \to 0} X^\varepsilon(x, y) = X(x, y)$
3. *$X^\varepsilon(x, y) > 0$ for some ε implies $\exists n \geq 0$ such that $0 < \lim_{\varepsilon \to 0} \varepsilon^{-n} X^\varepsilon(x, y) < \infty$*

Let $A_1, \ldots, A_m$ be the communicating classes of $X(M_1, M_2, \varepsilon)$. To leave a communicating class, there must be at least one incorrect signal.

Definition B.2.9 (Resistance)**.** *The resistance ρ_{ij} is the order of the probability that the system goes from communicating class A_i to A_j, i.e.*

$$\rho_{ij} = \min_{x \in A_i, y \in A_j, n \in \mathbb{N}} \mathcal{O}\left(X(M_1, M_2, \varepsilon)(x, y)^n\right),$$

where $\mathcal{O}(\cdot)$ is the order of the function. If the probability is 0, then the resistance is defined to be ∞.

Define the graph $\mathcal{G}$, which has one vertex, v_k, for every communicating class A_k. For every vertex pair, $v_i, v_j \in \mathcal{G}$, there is an edge with resistance ρ_{ij}.

Definition B.2.10 (i-tree)**.** *An i-tree in $\mathcal{G}$ is a spanning tree such that from every vertex $j \neq i$, there is a unique path directed from j to i.*

For each vertex, $\mathcal{T}_i$ is the set of all i-trees on $\mathcal{G}$. The resistance on an i-tree is,

$$\rho(\tau) = \sum_{(i,j)\in\tau} \rho_{ij}.$$

Definition B.2.11 (Stochastic Potential)**.** *The stochastic potential of the communicating class A_i is the least resistance among all i-trees:*

$$\gamma_i = \min_{\tau\in\mathcal{T}_i} \rho(\tau).$$

The stochastic potential measures the likelihood of the system visiting a certain communicating class. Communicating classes that don't have the minimum stochastic potential are at least an order ε less likely to be visited by the system. As the errors approach zero, the system spends non-trivial amounts of the supergame in only the communication classes with minimum stochastic potential. Finally, define the minimum stochastic potential of the system to be,

$$\gamma^* = \min_{i=1,\ldots m} \gamma_i.$$

Lemma B.2.12. *Given automata M_1 and M_2 subject to regular signal functions r_1 and r_2, the perturbed system $X(M_1, M_2, \varepsilon)$ is a regular perturbation.*

Proof

To show that this is true, I must show that the three criteria are satisfied. The system formed by automata M_1 and M_2 and regular signal functions r_1 and r_2 is represented by Markov chain $X(M_1, M_2, \varepsilon)$. By Lemma B.2.1, $X(M_1, M_2, \varepsilon)$ is always irreducible, so the first criterion is satisfied.

By the first part of the definition of regular signal function and (3.2), the second criterion is satisfied. Finally, it is clear that the third condition of the regular signal function remains under addition and multiplication, so the third criterion also holds by (3.2). ■

Theorem B.2.13 (Theorem 4 from Young (1993))**.** *Let X^0 be a stationary Markov process on a finite state space with communicating communication classes $A_1, \ldots, A_m$. Let X^ε be a regular perturbation of X^0, and let π^ε be its unique stationary distribution for every small positive ε. Then:*

1. *as $\varepsilon \to 0$, π^ε converges to a stationary distribution π^0 of X^0, and*

2. *x is stochastically stable ($\pi_x^0 > 0$) if and only if x is contained in a communicating class A_j that minimizes γ_j.*

The second part of this theorem implies that a communicating class is prevalent if and only if it minimizes stochastic potential.

Theorem 3.5.6. *Suppose players play supergame G with regular signal function r_i, and play automata $M_i \in \mathcal{M}_i^R$ represented by Markov chain $X(M_1, M_2, \varepsilon)$. If there exists some $\bar{\varepsilon} > 0$ such that (M_1, M_2) is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$, then for all prevalent communicating classes A_k, $U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i}))$.*

Proof of Theorem 3.5.6

Suppose that (M_1, M_2) is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$. Suppose by means of contradiction that there exists a communicating class A_k such that $\gamma(A_k) = \gamma^*$ and

$$U_2^{CC}(A_k) < U_2^{AC}(a_2^*(M_1)). \tag{B.7}$$

Using (B.7), Lemma B.2.5 and that fact that a communicating class can never get payoff higher than the optimal absorbing class gives,

$$U_2(M_1, M_2, \varepsilon) < U_2^{AC}(a_2^*(M_1)). \tag{B.8}$$

By Lemma 3.5.7, there exists an automaton M_2' such that for all communicating classes A of $X(M_1, M_2, \varepsilon)$, $U_2^{CC}(A) = U_2^{AC}(a_2^*(M_1))$. Therefore, by Lemma B.2.5,

$$\lim_{\varepsilon \to 0} U_2(M_1, M_2', \varepsilon) = U_2^{CC}(A) = U_2^{AC}(a_2^*(M_1)). \tag{B.9}$$

By Lemma 1 from Young (1993), the stationary distribution of $X(M_1, M_2', \varepsilon)$ is continuous at $\varepsilon = 0$. Therefore the payoff must also be continuous at $\varepsilon = 0$. So, for all $\varepsilon \in (0, \bar{\varepsilon})$, there exists some $\delta > 0$ such that,

$$\left| \lim_{\varepsilon \to 0} U_2(M_1, M_2', \varepsilon) - U_2(M_1, M_2', \varepsilon) \right| < \delta. \tag{B.10}$$

Set $\bar{\varepsilon}$ sufficiently small so that,

$$\left| \lim_{\varepsilon \to 0} U_2(M_1, M_2', \varepsilon) - U_2(M_1, M_2', \varepsilon) \right| < \left| U_2^{AC}(a_2^*(M_1)) - U_2(M_1, M_2, \varepsilon) \right|. \tag{B.11}$$

By (B.8), (B.9), and (B.11) for all $\varepsilon \in (0, \bar{\varepsilon})$,

$$U_2(M_1, M_2, \varepsilon) < U_2(M_1, M_2', \varepsilon).$$

So (M_1, M_2) is not an equilibrium for any $\varepsilon \in (0, \bar{\varepsilon})$, which is a contradiction. ■

Fix automaton M_1. This automaton has some optimal absorbing class $a_2^*(M_1)$. Let $\mathcal{M}^{SPM}(M_1)$ be the set of all automata $M_2 \in \mathcal{M}^R$ such that all communicating classes A_k of $X(M_1, M_2, \varepsilon)$ that minimize stochastic potential ($\gamma(A_k) = \gamma^*$) yield the optimal absorbing class payoff, $U_2^{CC}(A_k) = U_2^{AC}(a_2^*(M_1))$.

Theorem 3.5.8. *Suppose players play supergame G with regular signal function r_i, and play automata $M_i \in \mathcal{M}_i^R$ represented by Markov chain $X(M_1, M_2, \varepsilon)$. If*

1. *for all prevalent communicating classes A_k, $U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i}))$, and*
2. $\frac{\partial U_i(M_1, M_2, 0)}{\partial \varepsilon} = \sup_{M \in \mathcal{M}^{SPM}(M_{-i})} \frac{\partial U_i(M_i, M, 0)}{\partial \varepsilon}$;

then there exists some $\bar{\varepsilon} > 0$ such that (M_1, M_2) is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$.

Proof of Theorem 3.5.8

Fix (M_1, M_2) represented by $X(M_1, M_2, \varepsilon)$ such that for all stochastic potential minimizing communicating classes $\gamma(A_k) = \gamma^*$,

$$U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i})) \tag{B.12}$$

and

$$\frac{\partial U_i(M_1, M_2, \varepsilon)}{\partial \varepsilon} = \sup_{M \in \mathcal{M}^{SPM}(M_{-i})} \frac{\partial U_i(M_1, M, \varepsilon)}{\partial \varepsilon}.$$

Without loss of generality, I will show that when these conditions are satisfied, M_2 is a best response to M_1. For all $M_2 \notin \mathcal{M}^{SPM}(M_1)$, there exists a stochastic potential minimizing communicating class such that $U_2^{CC}(A_k) < U_2^{AC}(a^*(M_1))$. By Lemma B.2.5 and that fact that a communicating class can never get payoff higher than the optimal absorbing class,

$$U_2(M_1, M_2, \varepsilon) < U_2^{AC}(a_2^*(M_1)) \text{ for all } M_2 \notin \mathcal{M}^{SPM}(M_1). \tag{B.13}$$

For all $M_2 \in \mathcal{M}^{SPM}(M_1)$,

$$\lim_{\varepsilon \to 0} U_2(M_1, M_2, \varepsilon) = U_2^{AC}(a_2^*(M_1))$$

By continuity of U_2, this means that for all $\varepsilon \in (0, \bar{\varepsilon})$ for $\bar{\varepsilon}$ sufficiently small,

$$U_2(M_1, M, \varepsilon) < U_2(M_1, M', \varepsilon) \text{ for all } M \notin \mathcal{M}^{SPM}, M' \in \mathcal{M}^{SPM}.$$

So the best response to M_1 for $\varepsilon \in (0, \bar{\varepsilon})$ must come from the set $\mathcal{M}^{SPM}$. For all $M \in \mathcal{M}^{SPM}(M_1)$,

$$\lim_{\varepsilon \to 0} U_2(M_1, M, \varepsilon) = U_2^{AC}(a^*(M_1)).$$

By definition of the derivative, for some $\bar{\varepsilon} > 0$,

$$\left.\frac{\partial U_2\left(M_1, M, \varepsilon\right)}{\partial \varepsilon}\right|_{\varepsilon=0} \leq \left.\frac{\partial U_2\left(M_1, M', \varepsilon\right)}{\partial \varepsilon}\right|_{\varepsilon=0}$$

$$\Rightarrow U_2\left(M_1, M, \varepsilon\right) \leq U_2\left(M_1, M', \varepsilon\right) \text{ for all } \varepsilon \in (0, \bar{\varepsilon}) \ .$$

Therefore, if M_2 satisfies,

$$\frac{\partial U_i\left(M_1, M_2, \varepsilon\right)}{\partial \varepsilon} = \sup_{M \in \mathcal{M}^{SPM}(M_{-i})} \frac{\partial U_i\left(M_1, M, \varepsilon\right)}{\partial \varepsilon}.$$

Then it must be that for all $\varepsilon \in (0, \bar{\varepsilon})$,

$$U_2\left(M_1, M, \varepsilon\right) \leq U_2\left(M_1, M_2, \varepsilon\right) \text{ for all } M \in \mathcal{M}^{SPM}.$$

Therefore, M_2 is a best response to M_1. ■

B.2.2 Two-State Results

Proposition B.2.14. *Players play supergame G, where each action in stage game g has a unique best response. For any error $\varepsilon \in (0, 1/2]$, both players playing automata equivalent to open-loop finite automata is an equilibrium of the supergame G if and only if they play a Nash equilibrium of the stage game in every round of the supergame.*

If player 1 is playing a open-loop automaton, then it plays a fixed sequence of actions. The best response to this is simply to best respond in every round. If player 2's automaton is not equivalent to an open loop strategy and the chance of misperception is positive, then it is possible that a misperception could lead to a situation where player 2 doesn't best respond to player 1 in a given round.

If players play automata which always play the same action, and this action pair is a Nash equilibrium of the stage game, then this pair of automata always has to be a Nash equilibrium of the supergame. In addition, if the stage game has multiple Nash equilibrium, then any payoff in the convex hull of the Nash equilibrium payoffs is possible in equilibrium.

Proof of Proposition B.2.14

⇒First suppose that both players play automata equivalent to open loop automata M_1 and M_2. These form the Markov chain $X^{\varepsilon}_{M_1,M_2}$ with n states and all entries either 0 or 1. Depending on x^0, the Markov chain loops through $m \leq n$ states, $x^1, \ldots, x^m$. This yields payoff,

$$U_i\left(M_1, M_2, \varepsilon\right) = \frac{1}{m}\sum_{k=1}^{m} u_i\left(x^k\right).$$

Suppose without loss of generality that the actions in state x^j are not a Nash equilibrium of the stage game, because player 2 receives higher payoff from playing s_2^j than $f_2\left(q_2(x^j)\right)$ when player 1 plays $f_1\left(q_1(x^j)\right)$,

$$u_2\left(x^j\right) < u_2\left(f_1\left(q_1(x^j)\right), s_2^j\right).$$

Then player 2 is better playing automaton M' which is the same as M_2 except $f_2\left(q_2(x^j)\right) = s_2^j$,

$$U_2\left(M_1, M_2, \varepsilon\right) = \sum_{k\neq j} u_2(x^k) + u_i(x^j) < \sum_{k\neq j} u_2(x^k) + u_2(f_1\left(q_1(x^j)\right), s_2^j) = U_2\left(M_1, M', \varepsilon\right).$$

So both players playing automata equivalent to open loop automata M_1 and M_2 is an equilibrium only if a Nash equilibrium is played in every round.

⇐ Assume that automata M_1 and M_2 generate a sequence of actions which yield a Nash equilibrium in every stage game. Suppose that M_1 is not equivalent to an open loop automaton. For some state q_1,

$$f_1\left(\tau_1\left(q_1, C\right)\right) \neq f_1\left(\tau_1\left(q_1, D\right)\right).$$

So when M_1 is in q_1, the play in the next round can be either $f_1\left(\tau_1\left(q_1, C\right)\right)$ or $f_1\left(\tau_1\left(q_1, D\right)\right)$. Since $\varepsilon > 0$, either signal is possible with positive probability. Automaton M_2 will play s_2, which has a unique best response. So, with positive probability the system of automata M_1 and M_2 will not play a Nash equilibrium of the stage game. This contradicts the assumption that M_1 is not equivalent to an open loop automaton. A similar argument holds for M_2. Therefore, if M_1 and M_2 generate a sequence of action which yield a Nash equilibrium in every stage game that has unique best responses, the automata must be equivalent to open-loop automata. ■

Definition B.2.15 (Eventually Always Plays). *An automaton $M_i = (Q_i, q_i^0, f_i, \tau_i)$ eventually always plays action $s_i \in S_i$ if for all strongly connected components $Q_k^{SCC} \subseteq Q_i$,*

$$f_i(q) = s_i \text{ for all } q \in Q_k^{SCC}$$

Lemma B.2.16. *When $\varepsilon > 0$, and automaton M which eventually always plays C is payoff equivalent to M^C over the set of automata with only one SCC.*

Proof of Lemma B.2.16

Assume that player 1 plays $M_1 = M$. Assume that player 2 plays $M_2 = \left(Q_2, q_2^0, f_2, \tau_2\right)$ which has one strongly connected component. Let T_i be the round for which automata M_i reaches a strongly connected component. Since $\varepsilon > 0$, any sequence of signals occurs with positive probability, so $P(T_i < \infty) = 1$. Let u_i^t be the payoff for player i in round t. Let $T^* = \max(T_1, T_2)$. Then,

$$\begin{aligned} U_i(M_1, M_2, \varepsilon) &= \lim_{T\to\infty} \frac{1}{T}\left[\sum_{k=0}^{T^*} u_i^k + \sum_{k=T^*+1}^{T} u_i^k\right] \\ &= \lim_{T\to\infty} \frac{1}{T}\left[\sum_{k=T^*+1}^{T} u_i^k\right] \\ &= U_i(M', M_2, \varepsilon). \end{aligned}$$

So any automaton M with only one SCC is payoff equivalent to the automaton M' consisting only of the states of the SCC over the set of automata with only one strongly connected component. ■

Lemma B.2.17. *The set of two-state automata, $\mathcal{M}^2$, can be reduced to a smaller set of automata, $\bar{\mathcal{M}}^2$, such that,*

1. *for all* $M \in \mathcal{M}^2$*, there exists some* $M' \in \bar{\mathcal{M}}^2$ *such that* M *and* M' *are payoff equivalent over* $\mathcal{M}^2$*, and*

2. *for all* $M, M' \in \bar{\mathcal{M}}^2$*,* M *and* M' *are not payoff equivalent over* $\mathcal{M}^2$*.*

Proof of Lemma B.2.17

There are $|S_i|^N \left(N^N\right)^{|S_{-i}|}$ total N-state automata when the starting states are omitted. So when both players have two actions, there are 64 two-state automata. Many of these automata are redundant.

First, divide the 64 into four categories, each containing 16 automata:

$$\mathcal{M}_1^2 = \left\{M \in \mathcal{M}^2 | f(q_1) = C, f(q_2) = C\right\}$$

$$\mathcal{M}_2^2 = \left\{M \in \mathcal{M}^2 | f(q_1) = C, f(q_2) = D\right\}$$

$$\mathcal{M}_3^2 = \left\{M \in \mathcal{M}^2 | f(q_1) = D, f(q_2) = C\right\}$$

$$\mathcal{M}_4^2 = \left\{M \in \mathcal{M}^2 | f(q_1) = D, f(q_2) = D\right\}.$$

The automata in $\mathcal{M}_1^2$ play C regardless of the play of the other automaton. Therefore, these automata are equivalent to M^C, and hence payoff equivalent to M^C over the set $\mathcal{M}^2$. Similarly, the automata in $\mathcal{M}_4^2$ all play D regardless of the play of the other, so they are all payoff equivalent to M^D over $\mathcal{M}^2$.

For every $M_2 \in \mathcal{M}_2^2$, there exists an equivalent $M_3 \in \mathcal{M}_3^2$ (the only difference is that the states are switched). For example, $M_2 = (\{C, D\}, \{q_1, q_1\}, \{q_2, q_2\})$ and $M_3 = (\{D, C\}, \{q_2, q_2\}, \{q_1, q_1\})$. Both of these automata implement tit-for-tat, so they produce the same output regardless of the input, and hence are payoff equivalent. Without loss of generality, I only consider those automata in $\mathcal{M}_2^2$.

If automaton $M^E = (\{C, D\}, \{q_1, q_2\}, \{q_1, q_2\})$ starts in q_1, then regardless of the signals it plays C in every round of the supergame, and hence is equivalent to M^C. If M^E starts in q_2, then regardless of the signals, it plays D in every round, and hence is equivalent to M^D. So depending on the starting point, M^E is equivalent to either M^C or M^D. After equivalent automata have been eliminated, there are 17 remaining automata: M^C, M^D, and the set $\mathcal{M}_2^2 \backslash M^E$.

Note that all two-state automata have only one reachable SCC. For a two-state automaton to have multiple strongly connected components, each state needs to be a strongly connected component. The only two-state automaton which satisfies this is $M^E = (\{C, D\}, \{q_1, q_2\}, \{q_1, q_2\})$. If M^E starts in q_k, then only q_k can be reached, so it only has one reachable SCC, regardless of the starting point. Therefore, by Lemma B.2.16, any automaton which eventually always plays C is payoff equivalent to M^C over the set $\mathcal{M}^2$.

Out of the 17 remaining automata, three eventually always play C, and three eventually always play D ,

Eventually Always Play C	**Eventually Always Play D**
$(\{C, D\}, \{1,1\}, \{1,2\})$	$(\{C, D\}, \{1,2\}, \{2,2\})$
$(\{C, D\}, \{1,1\}, \{1,2\})$	$(\{C, D\}, \{2,2\}, \{1,2\})$
$(\{C, D\}, \{1,1\}, \{1,1\})$	$(\{C, D\}, \{2,2\}, \{2,2\})$

So by Lemma B.2.16, these automata are payoff equivalent to M^C and M^D over $\mathcal{M}^2$. The remaining 11 automata for the minimal set $\bar{\mathcal{M}}^2$.

1. M^C
2. M^D
3. M^{CD}
4. M^{WSLS}
5. $(\{C, D\}, \{1,1\}, \{2,1\})$
6. $(\{C, D\}, \{1,1\}, \{2,2\})$
7. $(\{C, D\}, \{2,1\}, \{1,1\})$
8. $(\{C, D\}, \{2,1\}, \{1,2\})$.
9. $(\{C, D\}, \{2,1\}, \{2,2\})$
10. $(\{C, D\}, \{2,2\}, \{1,1\})$
11. $(\{C, D\}, \{2,2\}, \{2,1\})$

■

Theorem 3.4.3. *In the infinitely repeated PD game, when players have the simple signal function r_i^S and choose among the set of two-state automata, $\mathcal{M}^2$, there are only three types of robust* *equilibria:*

1. $L < 0$ *and* M_i *is payoff equivalent to* M^C *for* $i = 1, 2$,

2. $L > 0$ *and* M_i *is payoff equivalent to* M^D *for* $i = 1, 2$, *and*

3. $-(1-2\varepsilon)^3 < L < (1-2\varepsilon)^3$ *and* $M_i = M^{WSLS}$ *for* $i = 1, 2$.

Proof of Theorem 3.4.3

If M_2 is the best response to M_1, then any automaton which is payoff equivalent to M_2 is also a best response to M_1. Therefore, I only need to consider the automata in the reduced payoff equivalent set $\bar{\mathcal{M}}(2)$ from Lemma B.2.17 when finding equilibria. However, if one of the automata in $\bar{\mathcal{M}}(2)$ is an equilibrium, then all payoff equivalent automata are also equilibria.

Three of the automata in $\bar{\mathcal{M}}(2)$ are open loop automata: M^D, M^C, M^{CD}. When $L \neq 0$, both players have unique best responses for all strategies in PD, so by Proposition B.2.14 these are equilibria if and only a Nash equilibrium is played in every stage game. Therefore, when $L > 0$, the unique Nash equilibrium of the stage game PD is for both players to play D. So M^D is an equilibrium when $L > 0$.

There are 8 remaining automata in $\bar{\mathcal{M}}^2$. For each of these automata M, I find the stationary distributions and payoffs when matched with each of the other automata in $\bar{\mathcal{M}}^2$. Using the payoffs, I calculate the best response function for each of the remaining 8 automata over almost all of the parameter space (all but set of measure zero). I find that the only regions which $M_1 = BR_1(M_2)$ and $M_2 = BR_2(M_1)$ are those stated in the theorem. For conciseness, these stationary distributions are not included here, but are available on my website.

The only equilibrium that is supported by a set of positive measure from these remaining 8 automata is M^{WSLS} on the region $-(1-2\varepsilon) < L < (1-2\varepsilon)$. So the three equilibria from $\bar{\mathcal{M}}^2$ are M^C, M^D, and M^{WSLS}.

There are also automata which are payoff equivalent to some of these three automata. By Lemma B.2.16, every automaton which eventually always plays C is payoff equivalent to M^C. Therefore, any combination of automata which eventually always play C is an equilibrium in the region $L < 0$. Similarly, any pair of automata which eventually play D is an equilibrium in the region $L > 0$. Finally, there are no other two-state automata which are payoff equivalent to M^{WSLS}. Therefore, the set of two-state equilibria that are supported by a region of positive measure is:

1. Both automata eventually always play C is a symmetric equilibrium if and only if $L > 0$.

2. Both automata eventually always play D is a symmetric equilibrium if and only if $L < 0$.

3. M^{WSLS} if and only if $-\left(1-2\varepsilon\right)^3 \leq L \leq \left(1-2\varepsilon\right)^3$.

■

	C	D
C	1+L,1	0,0
D	0,0	1,1+L

(a) BOS Game

	C	D
C	1+L,1+L	-L,0
D	0,-L	0,0

(b) MECG Game

Theorem B.2.18. *In the infinitely repeated BOS game, when players have the simple signal function r_i^S and choose among the set of two-state automata, $\mathcal{M}^2$, the only non-open-loop robust equilibria are:*

1. $-\frac{(1-2\varepsilon)^2}{2\varepsilon(2-5\varepsilon+4\varepsilon^2)} < L < \frac{(1-2\varepsilon)^2}{1-4\varepsilon+10\varepsilon^2-8\varepsilon^3}$ *and* $M_i = M^{WSLS}$.

Proof

The proof for this Theorem follows the argument of the proof of Theorem 3.4.3. Details available upon request. ■

Theorem B.2.19. *In the infinitely repeated MECG game, when players have the simple signal* *function* r_i^S *and choose among the set of two-state automata,* $\mathcal{M}^2$*, the only non-open-loop robust equilibria are:*

1. $L > \frac{1-4\varepsilon+10\varepsilon^2-8\varepsilon^3}{2\varepsilon(1-2\varepsilon)^2}$ *and* $M_i = M^{LSWS}$*, and*

2. $L > -\frac{1-8\varepsilon+14\varepsilon^2-8\varepsilon^3}{2(1-\varepsilon)(1-2\varepsilon)^2}$ *and* $M_i = M^{WSLS}$.

Proof

The proof for this Theorem follows the argument of the proof of Theorem 3.4.3. Details available upon request. ■

Theorem 3.4.4. *Suppose both players have simple signal functions* r_i^S*. If for* $i = 1, 2$*,*

1. $u_i(C,C) > u_i(C,D)$*, and*

2. $u_i(C,C) > \frac{u_i(D,C)+u_i(D,D)}{2}$*;*

then there exists some $\bar{\varepsilon} > 0$ *such that* $\left(M^{WSLS}, M^{WSLS}\right)$ *is an equilibrium for all* $\varepsilon \in (0, \bar{\varepsilon})$*.*

Proof of Theorem 3.4.4

To prove this theorem, I use the sufficient conditions for equilibria provided in Theorem 3.5.8. This says that to be an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$,

1. For all communicating classes such that $\gamma(A_k) = \gamma^*$, $U_i^{CC}(A_k) = U_i^{AC}(a^*(M_{-i}))$, and

2. $$\frac{\partial U_i(M_1, M_2, \varepsilon)}{\partial \varepsilon} = \sup_{M \in \mathcal{M}^{SPM}(M_{-i})} \frac{\partial U_i(M_1, M, \varepsilon)}{\partial \varepsilon}.$$

First assume that

$$u_i(C,C) > u_i(C,D) \tag{B.14}$$

and

$$u_i(C,C) > \frac{u_i(D,C) + u_i(D,D)}{2} \tag{B.15}$$

hold. I then show that the two sufficient conditions are satisfied, meaning $\left(M^{WSLS}, M^{WSLS}\right)$ is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$.

When both players play M^{WSLS}, the Markov chain for the system is,

$$X\left(M^{WSLS}, M^{WSLS}, \varepsilon\right) = \begin{bmatrix} (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 \\ (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & \varepsilon^2 \end{bmatrix}.$$

This system has one communicating class, A, consisting of the first state of the Markov chain. Since there is only one communicating class, it trivially minimizes stochastic potential. Therefore, it must be the case that the payoff in this communicating class is equal to the optimal absorbing class payoff. The payoff for the communicating class is,

$$U_i^{CC}(A) = u_i(C, C),$$

which is the stage-game payoff associated with joint action pair (C, C).

Next, I must calculate the optimal absorbing class payoff for M^{WSLS}. There are three extreme absorbing classes, such that any other absorbing class can be written as a convex combination of these extreme absorbing classes. So one of these has to be the optimal absorbing class.

1. $a_1(M^{WSLS}) = (\{q_1\}, \{C\})$ with payoff $u_i\left(a_1(M^{WSLS})\right) = u_i(C, C)$
2. $a_2(M^{WSLS}) = (\{q_1, q_2\}, \{D, D\})$ with payoff $u_i\left(a_2(M^{WSLS})\right) = \frac{u_i(D,C)+u_i(D,D)}{2}$
3. $a_3(M^{WSLS}) = (\{q_2\}, \{D\})$ with payoff $u_i\left(a_3(M^{WSLS})\right) = u_i(C, D)$

By (B.14) and (B.15), it is clear that $a_1\left(M^{WSLS}\right)$ is the optimal absorbing class. Therefore $U_i^{CC}(A) = U_i^{AC}\left(a_i^*\left(M^{WSLS}\right)\right)$, so the first condition is satisfied.

Next, I need to show that the marginal utility condition is satisfied. By Lemma B.2.17, the set of automata can be reduced some minimal payoff equivalent set. There are 11 remaining automata, call this set $\bar{\mathcal{M}}_2$. It can easily be verified that when M^{WSLS} is matched with any automaton $M \in \bar{M}_2$, then all communicating classes minimize stochastic potential.

There are only two automata, such that when paired with M^{WSLS}, all communicating classes yield the optimal absorbing class payoff. These are M^{WSLS} and $M^5 = (\{C, D\}, \{1, 1\}, \{2, 1\})$.

When both play M^{WSLS}, then the stationary distribution is,

$$\pi\left(M^{WSLS}, M^{WSLS}, \varepsilon\right) = \begin{bmatrix} 1 - 4\varepsilon + 7\varepsilon^2 - 4\varepsilon^3 \\ \varepsilon\left(1-\varepsilon\right) \\ \varepsilon\left(1-\varepsilon\right) \\ \varepsilon\left(2 - 5\varepsilon + 4\varepsilon^2\right) \end{bmatrix}'.$$

By Lemma 3.5.3, the payoff is the stationary distribution dotted with the vector of payoffs,

$$U_i\left(M^{WSLS}, M^{WSLS}, \varepsilon\right) = \pi\left(M^{WSLS}, M^{WSLS}, \varepsilon\right) \cdot \boldsymbol{u},$$

where $\boldsymbol{u}$ is the vector of payoffs. Therefore the marginal utility at $\varepsilon = 0$ is,

$$\frac{\partial U_i\left(M^{WSLS}, M^{WSLS}, 0\right)}{\partial \varepsilon} = -4u_i\left(C, C\right) + u_i\left(C, D\right) + u_i\left(D, C\right) + 2u_i\left(D, D\right)$$

When player 1 plays M^{WSLS} and player 2 plays M^5, then the stationary distribution is,

$$\pi\left(M^{WSLS}, M^5, \varepsilon\right) = \frac{1}{1 + 2\varepsilon - 6\varepsilon^2 + 10\varepsilon^3 - 4\varepsilon^4} \begin{bmatrix} 1 - 3\varepsilon + 5\varepsilon^2 - 2\varepsilon^3 \\ \varepsilon\left(1 - 2\varepsilon + 3\varepsilon^2 - 2\varepsilon^3\right) \\ \varepsilon\left(2 - 3\varepsilon + 2\varepsilon^2\right) \\ \varepsilon\left(2 - 6\varepsilon + 7\varepsilon^2 - 2\varepsilon^3\right) \end{bmatrix}.$$

Again by Lemma 3.5.3, the payoff is the dot product,

$$U_i\left(M^{WSLS}, M^5, \varepsilon\right) = \pi\left(M^{WSLS}, M^5, \varepsilon\right) \cdot \boldsymbol{u},$$

This means the marginal utility at $\varepsilon = 0$ is,

$$\frac{\partial U_i\left(M^{WSLS}, M^5, 0\right)}{\partial \varepsilon} = -5u_i\left(C,C\right) + u_i\left(C,D\right) + 2u_i\left(D,C\right) + 2u_i\left(D,D\right).$$

So

$$\frac{\partial U_i\left(M^{WSLS}, M^{WSLS}, 0\right)}{\partial \varepsilon} \geq \frac{\partial U_i\left(M^{WSLS}, M^5, 0\right)}{\partial \varepsilon} \iff u_i\left(C,C\right) > u_i\left(D,C\right).$$

This clearly holds by the assumption (B.14), and therefore both conditions are satisfied. So $\left(M^{WSLS}, M^{WSLS}\right)$ is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$ if the two conditions are satisfied. ■

B.3 Examples

B.3.1 Stochastic Potential Example

To better understand the definitions used for the theorem, I provide a corollary which shows that both players playing tit-for-tat can never be an equilibrium in the finite-state case. Let M^{TFT} be the two-state tit-for-tat automaton from Figure 3.6(a). Suppose players use the simple signal function r_i^S from (3.1). Finally suppose that players play supergame G with the prisoner's dilemma stage-game payoffs displayed in Figure 3.1.

Corollary B.3.1. *Suppose players play super game G with stage game PD and signal functions r_i^S, there is no $\bar{\varepsilon} > 0$ such that the pair of automata $\left(M^{TFT}, M^{TFT}\right)$ is an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$.*

Proof

To prove this, I need to show that the necessary conditions from Theorem 3.5.6 are not satisfied. The Markov chain of this system is,

$$X(M_1, M_2, \varepsilon) = \begin{matrix} x^{CC} \\ x^{CD} \\ x^{DC} \\ x^{DD} \end{matrix} \begin{bmatrix} (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ \varepsilon(1-\varepsilon) & \varepsilon^2 & (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) \\ \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 & \varepsilon^2 & \varepsilon(1-\varepsilon) \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 \end{bmatrix}.$$

There are three communicating classes: $A^C = \{x^{CC}\}, A^{CD} = \{x^{CD}, x^{DC}\}, A^D = \{x^{DD}\}$. The resistance matrix R which tells the resistance between each communicating class is,

$$R = \begin{matrix} A^C \\ A^{CD} \\ A^D \end{matrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}.$$

The entry in the first row, third column means that to probability of getting from A^C to A^D is order ε^2. The graph $\mathcal{G}$ with a vertex for each communicating class, and edge weights equal to the resistance between classes is displayed in Figure B.3.1(a). The optimal i-tree for each communicating class is displayed by the bold lines in Figure B.3.1(b)-(d). These graphs show that each communicating class has stochastic potential $\gamma_i = 2$. Therefore, the minimum stochastic potential for this system is $\gamma^* = 2$. By Theorem B.2.13, all communicating classes are prevalent. By Theorem 3.5.6, all prevalent communicating classes must yield the same payoff as the optimal absorbing class. The optimal absorbing class for each player yields payoff 1. Both players playing M^{TFT} only satisfies the necessary conditions if all communicating classes yield the same payoff, 1. Since $U_2(A^C) = 1$ and $U_2(A^D) = 0$, it is never possible for (M^{TFT}, M^{TFT}) to be an equilibrium for all $\varepsilon \in (0, \bar{\varepsilon})$. ■

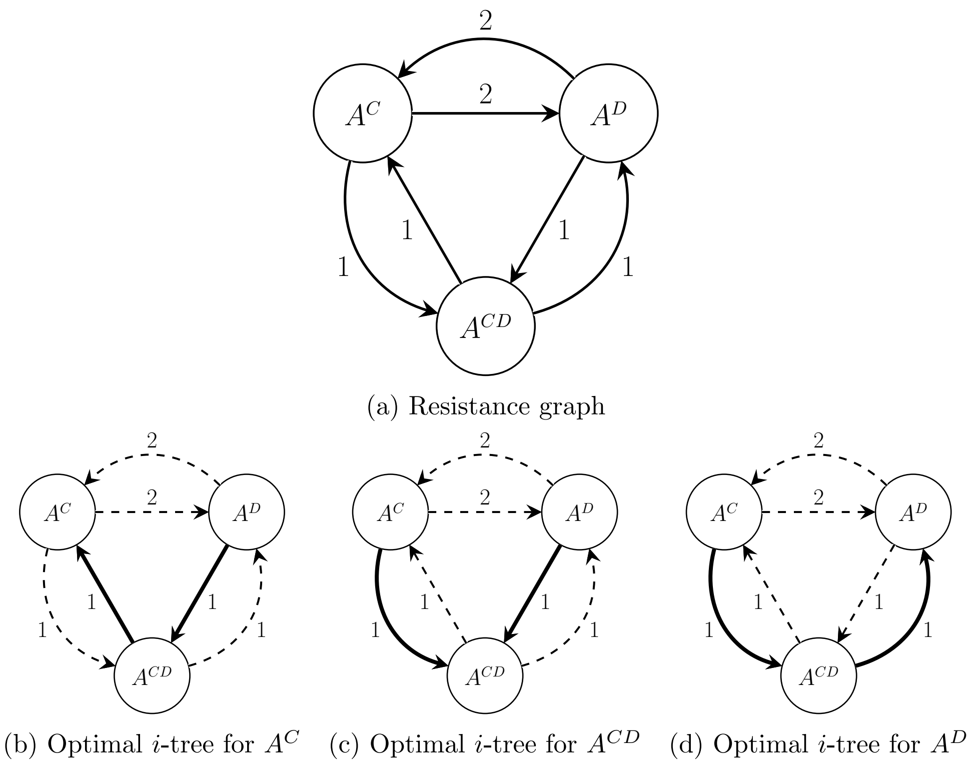


Figure B.3: Resistance graph and optimal i-trees if both players play M^{TFT}.

B.3.2 Constructed Automaton Example

Suppose that player 1 plays the three state automaton displayed in Figure B.4. First, player 2 wants to determine the desired absorbing class. For automaton M_1, the optimal absorbing class based on the prisoner's dilemma game from 3.1 is $a^*(M_1) = \{\{q_1\}, \{C\}\}$. Assume that player 2 wants to create an automaton M_2 which only gets stuck in this absorbing class. This automaton has three regions as described above, and is displayed in Figure B.5. First the absorbing region is simple, it

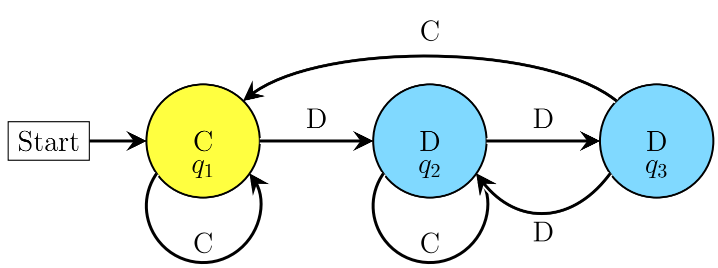


Figure B.4: Homing sequence example: automaton M_1.

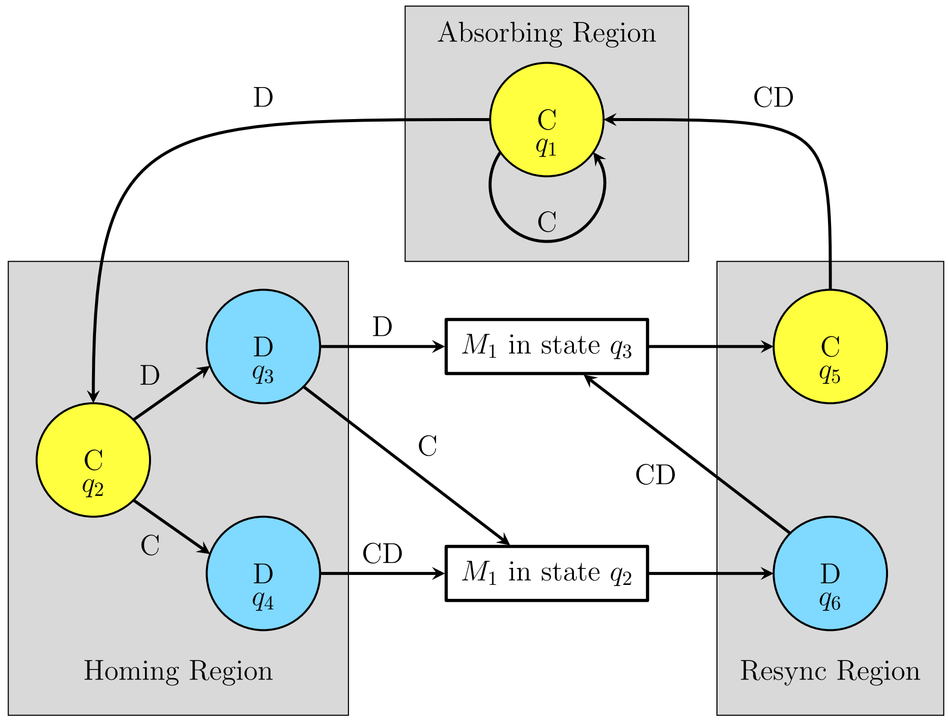


Figure B.5: Homing sequence example: constructed automaton M_2.

consists of one state, q_1, which plays C and returns when M_1 plays C. It is clear that when M_2 is in q_1, and M_1 is in q_1, player 2 is in his optimal absorbing class. If there is an incorrect signal while in the absorbing region, player 2 loses track of the current state of M_2, and therefore moves to the homing region to determine the current state.

The homing sequence for this automaton is $h = C, D$. To see why this is a homing sequence, suppose automaton M_1 starts in state q_1. Player 2 is trying to determine the current state by playing the homing sequence. In the first period, M_1 plays C and player 2 plays C. Automaton M_1 returns to state q_1. In the second period M_1 plays C again and player 2 play D. So the output from automaton M_1 from the homing sequence is C, C. The other sequences of plays for the other starting states is displayed in the following table:

Starting State	**First Play**	**Second Play**	**Final State**
q_1	C	C	q_2
q_2	D	D	q_3
q_3	D	C	q_2

When player 2 plays the homing sequence and sees output C, C or D, C, M_1 must be in state q_2. When the output is D, D, M_1 must be in q_3. So based on this output, player 2 knows the current state of M_1. The second region of M_2 is the homing region. In the homing region, M_2 always plays the homing sequence, and leaves the homing sequence after it has played this sequence. The homing region in Figure B.5 consists of states q_2, q_3, and q_4. In state q_2 the first term of the homing sequence is played, then depending on the output, M_2 moves to either state q_3 or q_4 where the second term of the homing sequence is played. The response from automaton M_1 after the homing region allows player 2 to know the current state of M_1. In this example, M_1 is either in state q_2 or q_3 after the homing region.

Finally, once the state has been determined, the automaton M_2 simply has to resynchronize the two automata back to the desired absorbing class $a_2^*(M_1)$. The resynchronization region consists of states q_5 and q_6. If M_1 is in state q_2, then automaton M_2 goes to state q_6. If automaton M_1 is in state q_3, then automaton M_2 goes to state q_5. After the resynchronization region, both automata are in state q_1, and they remain here until an incorrect signal is received.

Appendix C

Hysteresis in Coordination Games

C.1 Proofs

Lemma C.1.1. *In the game $cg(c)$ with $c \in (0,1)$, there are no asymmetric equilibria.*

Proof

Let σ_i be the probability that player i plays x_H. Let $\boldsymbol{\sigma} \in [0,1]^n$ be the vector of probabilities for all players. Suppose by means of contradiction that $\sigma_k \neq \sigma_j$. The payoffs for player i given that the other players play according to $\boldsymbol{\sigma}$ are

$$u_i\left(x_H, \boldsymbol{\sigma}_{-i}\right) = x_H \prod_{l \neq i} \sigma_l + x_L \left(1 - \prod_{l \neq i} \sigma_l\right) - c x_H, \text{ and}$$

$$u_i\left(x_L, \boldsymbol{p}_{-i}\right) = x_L \left(1 - c\right).$$

In order for $\boldsymbol{\sigma}$ to be a Nash equilibrium, both players must be indifferent between both of their actions, given the other players probabilities. So to be an equilibrium $\boldsymbol{\sigma}$ must satisfy,

$$x_H \prod_{l \neq i} \sigma_l + x_L \left(1 - \prod_{l \neq i} \sigma_l\right) - c x_H = x_L \left(1 - c\right) \text{ for all } i.$$

Since $x_L(1-c)$ is independent of $\boldsymbol{\sigma}$, then

$$x_H \prod_{l \neq i} \sigma_l + x_L \left(1 - \prod_{l \neq i} \sigma_l\right) - c x_H = x_H \prod_{l \neq j} \sigma_l + x_L \left(1 - \prod_{l \neq j} \sigma_l\right) - c x_H \text{ for all } i, j \in \mathcal{I}$$

which can be simplified to,

$$x_H \prod_{l \neq i} \sigma_l + x_L \left(1 - \prod_{l \neq i} \sigma_l\right) - x_H \prod_{l \neq j} \sigma_l - x_L \left(1 - \prod_{l \neq j} \sigma_l\right) = 0 \text{ for all } i, j \in \mathcal{I}$$

$$(x_H - x_L)(\sigma_j - \sigma_i) \prod_{l \neq i,j} \sigma_l = 0 \text{ for all } i, j \in \mathcal{I}.$$

The only way this equation can be satisfied for all $i, j \in \mathcal{I}$ is if $\sigma_i = \sigma_j$ for all i, j.

■

Proposition 4.2.3. *Using the symmetric quantal response equilibrium function f with parameter λ, the equilibrium correspondence $\boldsymbol{\Sigma}^*(c, \lambda)$ as λ varies has the following properties:*

1. $1/2 \in \boldsymbol{\Sigma}^*\left(\frac{1}{2^{N-1}}, \lambda\right)$ *for all* λ.

2. *The correspondence varies continuously from $\lambda = 0$ to $\lambda = \infty$ with endpoint $\boldsymbol{\sigma}_H = 1$ if $c < \frac{1}{2^{N-1}}$ and $\boldsymbol{\sigma}_H = 0$ if $c > \frac{1}{2^{N-1}}$.*

Proof

The first part follows from the functional form of (4.2). For the second part, rearrange (4.2) to get,

$$\lambda(\boldsymbol{\sigma}_H^*) = \frac{\ln\left(\frac{1-\boldsymbol{\sigma}_H^*}{\boldsymbol{\sigma}_H^*}\right)}{(x_H - x_L)\left({\boldsymbol{\sigma}_H^*}^{N-1} - c\right)}. \tag{C.1}$$

This is a well defined and continuous function for all $\boldsymbol{\sigma}_H^* \neq c^{\frac{1}{N-1}}$. The numerator of Equation (C.1) is a decreasing function of $\boldsymbol{\sigma}_H^*$, and is negative for $\boldsymbol{\sigma}_H^* > 1/2$ and positive for $\boldsymbol{\sigma}_H^* < 1/2$. The denominator is increasing in $\boldsymbol{\sigma}_H^*$ and changes signs when $\boldsymbol{\sigma}_H^* = c$. Therefore, the function $\lambda(\boldsymbol{\sigma}_H^*)$ is increasing for $c < {\boldsymbol{\sigma}_H^*}^{N-1}$ and decreasing for $c > {\boldsymbol{\sigma}_H^*}^{N-1}$.

It is clear that $\lambda\left(\frac{1}{2}\right) = 0$ for all values of c. Therefore if $\frac{d\lambda}{d\boldsymbol{\sigma}_H^*}\left(\frac{1}{2}\right) > 0$ then the correspondence varies continuously from $\lambda = 0$ to $\lambda = \infty$ with endpoint $\boldsymbol{\sigma}_H = 0$. Alternatively, if $\frac{d\lambda}{d\boldsymbol{\sigma}_H^*}\left(\frac{1}{2}\right) < 0$ then the correspondence varies continuously from $\lambda = 0$ to $\lambda = \infty$ with endpoint $\boldsymbol{\sigma}_H = 1$.

$$\frac{d\lambda}{d\boldsymbol{\sigma}_H^*}(\boldsymbol{\sigma}_H^*) = \frac{-\frac{(x_H - x_L)\left({\boldsymbol{\sigma}_H^*}^{N-1} - c\right)}{\boldsymbol{\sigma}_H^*(1-\boldsymbol{\sigma}_H^*)} - \ln\left(\frac{1-\boldsymbol{\sigma}_H^*}{\boldsymbol{\sigma}_H^*}\right)(N-1)(x_H - x_L)\left({\boldsymbol{\sigma}_H^*}^{N-2}\right)}{(x_H - x_L)^2\left({\boldsymbol{\sigma}_H^*}^{N-1} - c\right)^2}$$

$$\frac{d\lambda}{d\boldsymbol{\sigma}_H^*}\left(\frac{1}{2}\right) = \frac{-4}{(x_H - x_L)\left((1/2)^{N-1} - c\right)} \tag{C.2}$$

Since $x_H > x_L$, Equation (C.2) will be positive when $c > \frac{1}{2^{N-1}}$, and negative when $c < \frac{1}{2^{N-1}}$. ■

Proposition 4.2.4. *For every coordination game, $g(c) \in \mathcal{CG}$,*

1. *There exists a λ^* such that the logit SQRE correspondence, $\boldsymbol{\Sigma}^*(c, \lambda^*)$, exhibits hysteresis for all $\lambda > \lambda^*$, where,*
$$\lambda^* = \left(\frac{N}{N-1}\right)^N \frac{1}{x_H - x_l}.$$
2. *The critical value λ^* is decreasing in N.*
3. *For $N = 2$, the saddle-node bifurcation points are given by,*
$$\sigma_H = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4}{\lambda(x_H - x_L)}}.$$
4. *For $N > 2$, the saddle-node bifurcation points will not be symmetric around $\sigma_H = \frac{1}{2}$.*

Proof: First, calculate the symmetric quantal response equilibrium of the game. Suppose that all players play x_H with probability σ_H and x_L with probability σ_L. So the probabilities are,

$$P(\text{All others play H}) = \sigma_H^{N-1}, \text{ and}$$

$$P(\text{At least one other L}) = 1 - \sigma_H^{N-1}.$$

The payoffs are as follows,

$$
\begin{aligned}
u_i\left(x_L,\sigma_{-i}\right) &= \left(1-c\right)x_L \\
u_i\left(x_H,\sigma_{-i}\right) &= x_L(1-\sigma_H^{N-1}) + x_H\sigma_H^{N-1} - cx_H \\
&= x_L - cx_H + \sigma_H^{N-1}\left(x_H - x_L\right).
\end{aligned}
$$

Therefore, the symmetric logit quantal response equilibrium must satisfy the following equations,

$$
\begin{aligned}
\sigma_H &= \frac{e^{\lambda u_i(x_H,\sigma_{-i})}}{e^{\lambda u_i(x_L,\sigma_{-i})} + e^{\lambda u_i(x_H,\sigma_{-i})}} \\
&= \frac{1}{1+e^{\lambda[u_i(x_L,\sigma_{-i})-u_i(x_H,\sigma_{-i})]}} \\
&= \frac{1}{1+e^{\lambda(x_H-x_L)\left(c-\sigma_H^{N-1}\right)}}.
\end{aligned}
\tag{C.3}
$$

The symmetric QRE will be the σ_H^* that solves Equation (C.3). In order to show that hysteresis is possible, it is necessary to show that the bifurcation correspondence, $\mathbf{\Sigma}^*(c,\lambda)$, has the double saddle-node bifurcation. To do this, we find $c^*\left(\sigma_H\right)$, which is a function. Next, show that $\lim_{\sigma_H\to 0} c^*(\sigma_H) \to -\infty$, $\lim_{\sigma_H\to 1} c^*(\sigma_H) \to \infty$, and $c^{*\prime}(\sigma_H) < 0$ for some $\sigma_H \in (0,1)$. If these conditions hold then, the bifurcation correspondence, $\sigma_H^*(c)$, will have a double saddle-node bifurcation and look like an s-shaped curve. Rearrange Equation (C.3) to get,

$$
c^*(\sigma_H) = \sigma_H^{N-1} + \frac{\ln\frac{1-\sigma_H}{\sigma_H}}{\lambda\left(x_H - x_L\right)}. \tag{C.4}
$$

Which has a unique value for c^* for each value of σ_H. From this, notice that,

$$
\lim_{\sigma_H\to 0} c^*(\sigma_H) \to \frac{\ln\infty}{\lambda\left(x_H-x_L\right)} = \infty
$$

and,

$$
\lim_{\sigma_H\to 1} c^*(\sigma_H) \to 1 + \frac{\ln 0}{\lambda\left(x_H-x_L\right)} = \infty.
$$

Finally,

$$\frac{\partial c^*}{\partial \sigma_H}(\sigma_H) = (N-1)\,\sigma_H^{N-2} - \frac{1}{\lambda\,(x_H - x_L)}\left(\frac{1}{\sigma_H\,(1-\sigma_H)}\right). \tag{C.5}$$

Therefore,

$$\frac{\partial c^*}{\partial \sigma_H}(\sigma_H) > 0 \iff \lambda > \frac{1}{(N-1)\,p\,(1-p)^{N-1}\,(x_H - x_L)}.$$

In order to get the s-shaped curve, this needs to hold for some $\sigma_H \in (0,1)$. Since the right side of the above equation is minimized when $\sigma_H = {}^1\!/\!_N$, so see that,

$$\begin{aligned}
\frac{1}{(N-1)\,\sigma_H\,(1-\sigma_H)^{N-1}\,(x_H - x_L)} &\geq \frac{1}{(N-1)\,{}^1\!/\!_N\,(1-{}^1\!/\!_N)^{N-1}\,(x_H - x_L)} \\
&\geq \frac{1}{\frac{N-1}{N}^N\,(x_H - x_L)} \\
&= \left(\frac{N}{N-1}\right)^N \frac{1}{x_H - x_L}.
\end{aligned}$$

Therefore, if

$$\lambda^* = \left(\frac{N}{N-1}\right)^N \frac{1}{x_H - x_L},$$

then for all $\lambda \geq \lambda^*$, the bifurcation correspondence $\mathbf{\Sigma}^*(c,\lambda)(c)$ has the desired s-shaped form. Also, note that,

$$\left(\frac{N}{N-1}\right)^N \geq \left(\frac{N+1}{N}\right)^{N+1} \quad \text{for all } N \geq 2.$$

This holds by the Bernoulli Inequality[1]. Therefore, λ^* is decreasing in N. This means that as the

[1] The Bernoulli Inequality says that for $1 \geq \alpha > 0$ and $\delta \geq -1$,

$$(1+\delta)^\alpha \leq 1 + \alpha\delta.$$

Set $\alpha = {}^N\!/\!_{N+1}$ and $\delta = -\frac{1}{N}$, then the inequality tells us that,

$$\left(1 - \frac{1}{N}\right)^{N/N+1} \leq 1 - \frac{N}{N+1} \times \frac{1}{N} = \frac{N}{N+1}.$$

Taking the reciprocal,

$$\left(\frac{N}{N-1}\right)^{N/N+1} \geq \frac{N+1}{N}.$$

Or equivalently,

$$\left(\frac{N}{N-1}\right)^N \geq \left(\frac{N+1}{N}\right)^{N+1}.$$

group size increases, holding everything else constant, the s-shaped curve is more likely.

Finally for any fixed value of λ, the saddle-node bifurcation points of the s-shaped curve are at the two points where,

$$\frac{\partial c}{\partial \sigma_H}(\sigma_H) = 0.$$

Then set (C.5) to zero, and rearrange to get,

$$(1-\sigma_H)\,\sigma_H^{N-1} - \frac{1}{\lambda\,(x_H - x_L)\,(N-1)} = 0. \tag{C.6}$$

An explicit solution for this equation in not tractable unless $N = 2$. In the $N = 2$ case, solving this gives the solution

$$\begin{aligned} \sigma_H &= \frac{\lambda\,(x_H - x_L) \pm \sqrt{\lambda^2\,(x_H - x_L)^2 - 4\lambda\,(x_H - x_L)}}{2\lambda\,(x_H - x_L)} \\ &= \frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{4}{\lambda\,(x_H - x_L)}}. \end{aligned}$$

Also see that if $\lambda \geq \lambda^* = \frac{4}{x_H - x_L}$, then

$$1 \geq \frac{4}{\lambda\,(x_H - x_L)}.$$

So if $\lambda \geq \lambda^*$, then the two roots are always real, and if $\lambda < \lambda^*$, then there are no real roots, which is what we would expect. These two saddle-node bifurcation points are symmetric around $\sigma_H = \frac{1}{2}$ for the $N = 2$ case. However, for the $N > 2$ case, we would not expect to see this symmetry due to the form of (C.6).

Bibliography

ABREU, D., AND A. RUBINSTEIN (1988): "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56(6), 1259–1281.

ARIFOVIC, J., AND J. LEDYARD (2005): "Computer testbeds for mechanism design," *manuscript, April.*

AUMANN, R. (1981): *Essays in game theory and mathematical economics in honor of Oskar Morgenstern.* Bibliographisches Institut.

AXELROD, R. (1980a): "Effective Choice in the Prisoner's Dilemma," *The Journal of Conflict Resolution*, 24(1), 3–25.

——— (1980b): "More Effective Choice in the Prisoner's Dilemma," *The Journal of Conflict Resolution*, 24(3), 379–403.

——— (2006): *The evolution of cooperation.* Basic Books, revised edn.

BALDWIN, R. (1988): "Hysteresis in Import Prices: The Beachhead Effect," *The American Economic Review*, 78(4), 773–785.

BERGER, U. (2007): "Two more classes of games with the continuous-time fictitious play property," *Games and Economic Behavior*, 60(2), 247 – 261.

BHASKAR, V. (2000): "The robustness of repeated game equilibria to incomplete payoff information," *University of Essex.*

BHASKAR, V., AND I. OBARA (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40 – 69.

Blanchard, O. J., and L. H. Summers (1986): "Hysteresis and the European Unemployment Problem," *NBER Macroeconomics Annual*, 1, 15–78.

Brandts, J., and D. J. Cooper (2006): "A Change Would Do You Good .... An Experimental Study on How to Overcome Coordination Failure in Organizations," *The American Economic Review*, 96(3), 669–693.

Brown, G. (1951): "Iterative solution of games by fictitious play," *Activity analysis of production and allocation*, 13, 374–376.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A Cognitive Hierarchy Model of Games*," *Quarterly Journal of Economics*, 119(3), 861–898.

Carlsson, H., and E. v. Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.

Cole, H. L., and N. R. Kocherlakota (2005): "Finite memory and imperfect monitoring," *Games and Economic Behavior*, 53(1), 59 – 72.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66(3), 597–626.

Compte, O., and A. Postlewaite (2009): "Effecting Cooperation," .

Conlisk, J. (1996): "Why Bounded Rationality?," *Journal of Economic Literature*, 34(2), 669–700.

Cooper, R. W., D. V. DeJong, R. Forsythe, and T. W. Ross (1990): "Selection Criteria in Coordination Games: Some Experimental Results," *The American Economic Review*, 80(1), 218–233.

Dal Bó, P., and G. R. Fréchette (2008): "The evolution of cooperation in infinitely repeated games: Experimental evidence," *Unpublished manuscript. Brown University and New ....*

Dasarathy, B. (1991): *Nearest neighbor (NN) norms: NN pattern classification techniques*. Los Alamitos, CA: IEEE Computer Society Press.

David, P. A. (1985): "Clio and the Economics of QWERTY," *The American Economic Review*, 75(2), 332–337.

Dixit, A. (1989): "Hysteresis, Import Penetration, and Exchange Rate Pass-Through," *The Quarterly Journal of Economics*, 104(2), 205–228.

Ely, J. C., J. Horner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2), 377–415.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Friedman, D. (1991): "Evolutionary Games in Economics," *Econometrica*, 59(3), 637–666.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997–1039.

Goeree, J. K., and C. A. Holt (2005): "An experimental study of costly coordination," *Games and Economic Behavior*, 51(2), 349 – 364, Special Issue in Honor of Richard D. McKelvey.

Grinstead, C. M., and J. L. Snell (1997): *Introduction to probability*. American Mathematical Society, second revised edn.

Hahn, S. (2008): "The convergence of fictitious play in games with strategic complementarities," *Economics Letters*, 99(2), 304 – 306.

Harsanyi, J. C., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*, vol. 1 of *MIT Press Books*. The MIT Press.

Hart, S. (2002): "Evolutionary dynamics and backward induction," *Games and Economic Behavior*, 41(2), 227 – 264.

Imhof, L. A., D. Fudenberg, and M. A. Nowak (2007): "Tit-for-tat or win-stay, lose-shift?," *Journal of Theoretical Biology*, 247(3), 574 – 580.

Ioannou, C. (2009): "Algorithmic Bounded Rationality, Noise And Optimality," .

Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102(1), 1 – 15.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61(1), 29–56.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66(3), 627–652.

Knez, M., and C. Camerer (1994): "Creating Expectational Assets in the Laboratory: Coordination in 'Weakest- Link' Games," *Strategic Management Journal*, 15, 101–119.

Kohavi, Z. (1978): *Switching and Finite Automata Thoery*. McGraw-Hill Inc., 2nd edn.

Lambson, V. E., and D. A. Probst (2004): "Learning by matching patterns," *Games and Economic Behavior*, 46(2), 398 – 409.

Lehrer, E. (1988): "Repeated games with stationary bounded recall strategies," *Journal of Economic Theory*, 46(1), 130 – 144.

Mailath, G., and L. Samuelson (2006): *Repeated games and reputations: long-run relationships*. Oxford University Press, USA.

Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102(1), 189 – 228.

McKelvey, R., and T. Palfrey (2001): "Playing in the dark: Information, learning and coordination in repeated games," Discussion paper, Technical report, California Institute of Technology, 2001.

McKelvey, R. D., and T. R. Palfrey (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6 – 38.

Milgrom, P., and J. Roberts (1991): "Adaptive and sophisticated learning in normal form games," *Games and Economic Behavior*, 3(1), 82 – 100.

Miller, J. (1996): "The coevolution of automata in the repeated Prisoner's Dilemma," *Journal of Economic Behavior & Organization*.

Miyasawa, K. (1961): "On the convergence of the learning process in a 2 x 2 non-zero-sum two-person game," *Economic Research Program, Research Memorandum*, 33.

Monderer, D., and L. S. Shapley (1996): "Fictitious Play Property for Games with Identical Interests," *Journal of Economic Theory*, 68(1), 258 – 265.

Neyman, A. (1985): "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Economics Letters*, 19, 227–229.

Nowak, M., and K. Sigmund (1993): "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game," *Nature*, 364(6432), 56–58.

Nyberg, S. (1997): "The honest society: Stability and policy considerations," *Journal of Public Economics*, 64(1), 83 – 99.

Obara, I. (2009): "Folk theorem with communication," *Journal of Economic Theory*, 144(1), 120 – 134.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70 – 83.

Rankin, F. W., J. B. Van Huyck, and R. C. Battalio (2000): "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games," *Games and Economic Behavior*, 32(2), 315 – 337.

Rapoport, A., M. Guyer, and D. Gordon (1976): *The 2 x 2 game*. Univ of Michigan Pr.

Rubinstein, A. (1986): "Finite Automata Play the Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 39, 83–96.

Sabourian, H. (1998): "Repeated games with M-period bounded memory (pure strategies)," *Journal of Mathematical Economics*, 30(1), 1 – 35.

Selten, R., and T. Chmura (2008): "Stationary Concepts for Experimental 2x2-Games," *The American Economic Review*, 98(3).

Simon, H. (1957): *Models of man: Social and rational.* John Wiley.

Smith, J. M., and G. R. Price (1973): "The Logic of Animal Conflict," *Nature*, 246(5427), 15–18.

Sonsino, D. (1997): "Learning to Learn, Pattern Recognition, and Nash Equilibrium," *Games and Economic Behavior*, 18(2), 286 – 331.

Sonsino, D., and J. Sirota (2003): "Strategic pattern recognition–experimental evidence," *Games and Economic Behavior*, 44(2), 390 – 411.

Stahl, D., and J. B. Van Huyck (2002): "Learning conditional behavior in similar stag hunt games," .

Stigler, G. J. (1964): "A Theory of Oligopoly," *The Journal of Political Economy*, 72(1), 44–61.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *The American Economic Review*, 80(1), 234–248.

Weber, R. A. (2006): "Managing Growth to Achieve Efficient Coordination in Large Groups," *The American Economic Review*, 96(1), 114–126.

Wedekind, C., and M. Milinski (1996): "Human Cooperation in the Simultaneous and the Alternating Prisoner's Dilemma: Pavlov Versus Generous Tit-For-Tat," *Proceedings of the National Academy of Sciences of the United States of America*, 93(7), 2686–2689.

Yamamoto, Y. (2009): "A limit characterization of belief-free equilibrium payoffs in repeated games," *Journal of Economic Theory*, 144(2), 802 – 824.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.