Applications of Individual Evolutionary Learning

by

Joshua Forrest Boitnott

M.A., Simon Fraser University, 2009
B.A., Simon Fraser University, 2008

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Economics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY
Spring 2017

Approval

Name: **Joshua Forrest Boitnott**

Degree: **Doctor of Philosophy (Economics)**

Title: ***Applications of Individual Evolutionary Learning***

Examining Committee: **Chair:** Dr. Simon Woodcock, Associate Professor
Department of Economics

Dr. Jasmina Arifovic
Senior Supervisor
Professor

Dr. Erik Kimbrough
Supervisor
Associate Professor

Dr. John Duffy
Supervisor
Professor
Department of Economics
University of California, Irvine

Dr. Steeve Mongrain
Internal Examiner
Professor

Dr. Nobuyuki Hanaki
External Examiner
Professor
Department of Economics
University of Nice Sophia Antipolis

Date Defended: 30 January 2017

Abstract

This research investigates three applications of the Individual Evolutionary Learning (IEL) model. Chapter 2 utilizes a horse-race approach to investigate the overall performance of 4 learning algorithms in games with congestion. The games utilized are Market Entry games and Choice of Route games. I show that a version of the IEL has the best fit of the experimental data relative when the experimental subjects have full information. Chapter 3 (joint work the Jasmina Arifovic and John Duffy) applies the IEL to games with correlated equilibrium suggested by an external third party. The IEL nearly perfectly matches the behavior of experimental subjects playing the Battle of the Sexes game, but requires an adjustment to the initial conditions to match the behavior of experimental subjects in the Chicken game. Chapter 4 extends the Individual Evolutionary Learning with Other-Regarding Preferences (IELORP*) model to force the algorithm to match the discrete nature of the experimental choices and introduce beliefs via adaptive expectations. The algorithm continues to match the stylized facts associated with the standard LPGG, but does not appear to extend to games where beliefs are elicited using monetary incentives.

Keywords: Agent-Based Modeling, Evolutionary Algorithm, Experimental Economics, Learning

Dedication

To Ana: My Dearest Love, Wife, and Partner

Acknowledgements

I would like to begin by thanking my family for their patience as I have pursued my doctoral thesis. I know that I would have been unable to finish without their love and support. I need to give a great deal of gratitude to my Senior Advisor Jasmina Arifovic who has supported my research agenda. I offer my thanks to Erik Kimbrough for providing excellent feedback that has been instrumental in improving several of my papers. I would like to thank John Duffy for taking the time to work with me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 1 serves as an introduction and provides a brief overview of each chapter. All of the chapters in this thesis focus on different applications and extensions of the Individual Evolutionary Learning (or IEL) model. The Individual Evolutionary Learning model is a trial and error learning model that seeks to match behavioral regularities observed in experimental settings. One of the appealing aspects of the IEL is its use in modeling a wide variety of experiments while maintaining a fairly consistent set of parameter values. This consistency suggests a level of transferability not applicable to many other learning algorithms where the parameter values need to be re-estimated to model human behavior in experimental settings.

Chapter 2 focuses on the behavior of four learning algorithms in games that display congestion. Learning algorithms represent an important step in modeling the dynamic behavior of human subjects observed in experimental settings, yet many existing algorithms have limited application outside of a single game or across different experimental settings when parameter values are fixed. Furthermore, while previous studies have investigated algorithmic agent performance relative to experimental data, they have relied upon data from a single source. In this study, I use a horse-race approach to compare the performance of four learning algorithms (Impulse-Matching Learning, Self-Tuning Experience Weighted Attractions, and Individual Evolutionary Learning(x2)) across five different sets of experimental data generated by multiple researchers where the subjects face congestion. I examine the ability of each algorithm to replicate experimental data independent of the training data set used to determine the initial parameter values via the root mean squared error for each algorithm relative to the data produced by the human subjects and the ability of the algorithms to match multiple experimental equilibria despite adjusting the payoffs, the number of players, and the number of repetitions played. The Individual Evolutionary Learning model with a modified set of assumptions is determined to have the closest fit to the regularities seen in the experimental data when human subjects were given full information measured by the average mean squared distance, but not in experiments with limited information.

Chapter 3 is joint work with Jasmina Arifovic and John Duffy examining the behavior of the IEL relative to human subject data from games with correlated equilibrium. Correlated equilibrium (Aumann 1974, 1987) is an important generalization of the Nash equilibrium concept for multiplayer non-cooperative games. In a correlated equilibrium, players rationally condition their strategies on realizations of a common external randomization device and, as a consequence, can achieve payoffs that Pareto dominate any of the game's Nash equilibria. In this chapter, we explore whether such correlated equilibria can be *learned* over time using an evolutionary learning model where agents do not start with any knowledge of the distribution of random draws made by the external randomization device. Furthermore, we validate our learning algorithm findings by comparing the end behavior of simulations of our algorithm with both the correlated equilibrium of the game and the behavior of human subjects that play the same game. Our results suggest that the evolutionary learning model is capable of learning the correlated equilibria of these games in a manner that approximates well the learning behavior of human subjects and that our findings are robust to changes in the specification and parameterization of the model.

Chapter 4 presents an extension of the Individual Evolutionary Learning with Other-Regarding Preferences (IELORP*) model be Arifovic & Ledyard (2012). The extension focuses on two major components: matching the experimental setting and introducing a system of non-myopic beliefs. To match the linear public goods experimental setting as closely as possible, the set of choices available to the agents in the model are adjusted to match the discrete actions available in experimental settings. The beliefs used in the model are adaptive expectations. The introduction of beliefs allows for a more sophisticated version of the model that is able to have a different initialization process. Once the model is described, I show the model's simulated agents continue to match experimental data from the standard Linear Public Goods Game. The model is then applied to linear public goods experiments where beliefs about the actions of others are elicited via a monetary payoff using a proper scoring rule. The simulations are generally unable to match the experimental data as well as the experiments from the standard Linear Public Goods Game, but this result may be the result of differences in experimental design or in the scoring rule used.

Writing this thesis has allowed me the freedom and time to explore many of the intricacies of the Individual Evolutionary Learning model. In the process, the relative parsimony of the model has become more apparent as the general model is able to be applied across many different games. The model is therefor easy to transfer into many different settings. The major alteration necessary for successful implementation of the model is determining the shape and nature of the strategy to be used. For instance, the probabilistic strategy used by the IEL in Chapter 2 would be insufficient to address the concept of correlated equilibrium examined in Chapter 3; however, the strategy choice could be readily applied to similar games such as the El Farol Bar Problem. Additionally, the IEL model in Chapter 2 found the incomplete information of the Erev et al (2010a) games difficult to match. This is unsurprising at the IEL has long had problems dealing with incomplete information games, but the introduction of a longer history as done in Chapter 3 might have improved the overall performance. Finally, the introduction of beliefs though unsuccessful in matching the behavior of experimental subjects in incentivized games provides a baseline for understanding the overall importance of a relatively more sophisticated belief system. The beliefs developed in Chapter 4 could be applied to the games examined in Chapter 2 since these games rely on individuals adjusting their actions and beliefs over time.

Chapter 5 concludes this thesis. The discussion focuses the overall knowledge gained from this thesis along with suggestions for future areas of research.

Chapter 2

The Behavior of Learning Algorithms in Experimental Games with Congestion

2.1 Introduction

Congestion is an oft-unavoidable consequence of scarcity and can have significant negative consequences on nearly everything in life from transportation and urban planning to the enjoyment of YouTube videos. Analysis of congestion typically utilizes a static model of behavior to predict the equilibrium that does not match the dynamic human subjects behavior observed in experimental settings. Developing an accurate model of human behavior in markets with congestion would allow for more efficient planning. In order to choose the best model, it is important to understand how well existing learning models capture behavior in these games. By programming a set of algorithms to play games and comparing with human subject data, I can determine if an existing or modified learning algorithm is able to mimic the human behavior displayed in experimental settings.

Congestion is a negative externality imposed on an individual when other individuals make the same choice. Market entry games and route choice games represent two examples. A market entry game has an agent decide whether to enter a market (or not) with congestion only influencing the payoff of one choice (not the other). The classic example of a congestion game is the route choice game where players in a network must choose between two or more routes, and the more players that are on any particular route the lower is the overall payoff for each agent selecting that route. While there are a plethora of games where crowding exists, this chapter limits itself to the market entry game and the route choice game since these games capture the strategic uncertainty associated with congestion and apply across multiple experimental settings.

Modeling behavior in congestion games has a long history in economics with Rosenthal (1973) demonstrating there always exists at least one pure strategy Nash equilibrium (PSNE) in a general model. Often a multitude of equilibria exist in these games including symmetric mixed strategy Nash equilibrium (MSNE). Because of the importance of congestion and variety of possible equilibria, experimental economists have taken up investigating how agents behave in markets with congestion to find what equilibrium predictions match.[1]

In market entry games, the externality is limited to a single action (i.e. enter the market) with the alternative action providing a payoff independent of the actions taken by others. Experimental economists have generally found subjects choose to enter the market more frequently than would be supported by the option without the congestion. Both Duffy and Hopkins (2005) and Erev et al (2010a) reported experimental subjects playing enter more often than was predicted by the PSNE or the MSNE with the excessive entry declining over time. The distribution predicted by the PSNE was eventually matched by the human subjects in Duffy and Hopkins, but took 90 periods to reach this outcome.

Unlike the market entry game, the route choice game features a congestion externality regardless of the action taken. The experimental evidence generally supports the idea that the mean route choice is well described by the distribution predicted by either the MSNE or the PSNE. But the experimental data shows subjects exhibit persistent fluctuations that do not match the predictions of the PSNE or the MSNE. Despite the different payoff functions used, both Chmura and Pitz (2006) and Rapoport et al (2009) found aggregate route choice consistent with the MSNE when players were facing symmetric payoff functions. Selten et al (2007) used an asymmetric payoff function that produced average route choice consistent with the PSNE distribution. Additionally, a different portion of the experiment in Rapoport et al also found support for players moving toward the PSNE when the payoff functions were asymmetric.

In order to allow human subjects time to learn to play, all of the experiments discussed allowed for repeated interactions by the same group of people, thus creating a dynamic process. Because I am interested in mimicking the process of discovery displayed by the experimental subjects, a set of learning algorithms have been applied to these same games. The use of artificial agents has a long history in economics with a recent trend to develop more "human like" artificial constructs using learning algorithms to further our understanding of behavior in experimental settings[2]. Recent examples of papers trying to find "human like" agents have implemented "horse-races" to gain a better understanding of what "agent-based modeling approaches" (Duffy, 2006 p. 1004) should be used. Erev et al implemented a horse-race via a tournament for algorithmic learning models using data collected from a market entry game. Chmura et al (2012) utilized a horse-race approach to examine the

[1]This chapter neglects the experimental literature discussing how to achieve the efficient outcome as not all markets that are subject to congestion allow for clear and easily identifiable congestion pricing solutions.

[2]See Chen (2012) for an excellent overview of agent-based computational economics (ACE) with the proposal to find artificial agents that could mimic human behavior discussed in Arthur (1993).

performance of several new learning algorithms relative to a set of commonly used learning algorithms in 2x2 games. Both Erev et al and Chmura et al attempt to address some issues that have plagued the application of algorithms in economic settings by using larger data sets, trying to avoid over-fitting the data, and considering a number of different algorithms; however, neither paper moves to the next step of applying the models to different experimental data. In this chapter, a set of four algorithms will be applied to each game, roughly fitting parameter values or using previously successful parameters or implementing a model without parameter values, and comparing the outcomes with data produced in the different experiments. Unlike the previously mentioned papers utilizing a horse-race approach that only varied the payoff function, the different experiments being examined here vary in the number of agents participating, the number of choices available, and the total number of periods played in the experiments. This study attempts to find an algorithm that matches the different experimental behavior displayed across an array of adjustments.

The learning algorithms studied in this chapter are Impulse-Matching Learning, Self-Tuning Experience Weighted Attraction, Individual Evolutionary Learning, and an Individual Evolutionary Learning Alternative. The success of the Impulse-Matching Learning and Self-Tuning Experience Weighted Attraction in Chmura et al supports the inclusion. The Impulse-Matching Learning is based on people responding positively to outcomes that achieved high payoffs and negatively to outcomes that achieved low payoffs (from Learning Direction Theory (Selten et al, 2005)) with losses and gains weighted differently (from Prospect Theory (Kahneman and Tversky, 1979)) based on an aspiration level equal to the maximin payoff. To accomplish the task of weighting loss and gains differently, the payoff function of the game is transformed to create impulses that reinforce the choice with the highest payoff. Thus Impulse-Matching Learning appears to place a large demand on the amount of knowledge necessary for the algorithm to operate properly. The Self-Tuning Experience Weighted Attraction uses Learning Direction Theory to simplify the Experience Weighted Attraction by Camerer and Ho (1999) and minimizes the number of exogenous parameters needed. The Individual Evolutionary Learning has also successfully mimicked the data produced by experimental subjects in the linear public goods game (Arifovic and Ledyard, 2012) and correlated equilibria experiments (Arifovic et al, 2015). The model relies on the evaluation of hypothetical play by agents where strategies that perform well are reinforced, strategies that perform poorly are eliminated, and new possible strategies are introduced over time. Donoso et al (2014) show a similar process where the introduction of new ideas and evaluation of old ones appears to occur in the prefrontal cortex; however, the paper suggests people rely on fewer strategies than the Individual Evolutionary Learning model traditionally used. The Individual Evolutionary Learning Alternative is introduced to demonstrate how access to fewer strategies can improve the performance of the Individual Evolutionary Learning model here.

The approach in this chapter is to use the experimental data reported from a number of different experiments to examine what algorithms are able to produce outcomes consistent with the results across different settings where congestion externalities exist. While all of the experiments focus on the aggregate behavior in games with congestion, each experiment presents a unique focus giving rise to different features being measured and reported. This leads to the question: how can or should the experimental data be compared to the behavior of the algorithms? This chapter reports how the algorithms perform relative to the aspects reported in each paper via the root mean squared error and then examines the overall reliability of the algorithm across the games. The root mean squared error measures how well the individual runs of each algorithm are able to capture the behavior exhibited in the data. To examine the overall performance of the algorithms, I begin by testing if the results produced by the algorithms are statistically different from the predictions of the MSNE and PSNE. To make a reasonable comparison across games, a measure creating a single value needs to be implemented despite the different number of players, repetitions, and choices available in the various games. While several measures have been proposed and used such as mean squared deviation (see Feltovich (2000) or Chen and Khoroshilov (2003)), the proportion of inaccuracy (see Erev and Roth (1998) or Feltovich (2000)), the normalized mean squared deviations (see Erev et al), or the mean quadratic distance (see Chmura et al), these measures have generally been applied to cases where the number of actions or number of players is consistent over time. This chapter reports the average sum of squared distance created from the proportion or frequency each action is played, a measure similar to the mean quadratic distance. This gives another baseline for comparing the performance of the algorithms across the games despite the difference in number of agents and actions available.

The chapter begins with a general discussion of the games used including a review of the similarities and differences between the experiments. The next section gives a general description of the algorithms implemented. The final section provides a greater level of detail about the games along with any adjustments needed for the algorithms to work. The root mean squared error is reported. Next, the chapter shows whether the algorithm simply reproduces the MSNE or PSNE before concluding with the average sum of squared distance. On average, all of the algorithms perform adequately if the focus is the average attendance (entry) in full information games where agents only have two actions. In more complex games, some of the algorithms fail to match experimental data. Appendix A.1 details two additional algorithms implemented: Impulse Balance Learning and Experience Weight Attraction with Appendix A.4 providing more tables, graphs, and information about the performance of all of the algorithms in each of the papers. In general, there is little discernible difference between the Impulse Balance Learning and the Impulse-Matching Learning; although the Impulse-Matching Learning clearly outperforms the Impulse Balance Learning when more than two actions are available to the agent. The parameter values for

Experience Weighted Attraction model are selected using a grid search of data from Duffy and Hopkins, yet the algorithm produces the highest root mean squared errors for several games.

2.2 The Games

In this section, a general overview of the games is given along with a discussion of the data and conclusions produced by the experiments. Each of the experiments was repeatedly played by a set of subjects grouped together for the entire experiment. The number of subjects varied from 4 to 18, the number of repetitions (indicated Rep in Table 2.1) varied from 40 to 200 periods of play, and the number of observations (indicated Obs in Table2.1) varied from 3 to 9. Table 2.1 contains a summary of the information included in this section. All of the games provided subjects with full information about the game except for Erev et al. A more detailed description of each game is given in Section 4 when the root mean squared errors are reported. First the payoff functions are discussed, followed by an overview of the predictions and outcomes produced in the experiments. In this section and throughout the rest of the tables in this chapter, abbreviations are used to identify the games from certain papers: D & H for Duffy and Hopkins (2005), C & P for Chmura and Pitz (2006), Selten for Selten et al (2007), Rapoport for Rapoport et al (2009), Rapoport 1 for the first set of route choice games in Rapoport et al, Rapoport 2 for the second set of route choice games in Rapoport et al, and Erev for Erev et al (2010). I finish by describing the test used to determine if an iteration (or run)[3] of the algorithm is different from the MSNE and PSNE predictions.

Despite the use of two different types of games, there are a number of similarities and differences in the payoff functions used. All of the games examined here capture congestion in the form of a constant cost function $c(m_j)$ where $c'(m_j) \geq 0$, $c''(m_j) = 0$, and $m_j \in N$ is the total number of agents choosing action or route j. The market entry (indicated ME in Table 2.1) games have a payoff for entering the market of $\pi = E - c(m_E)$ and a payoff for not entering the market of $\pi = v$ for $E - c(0) > v > E - c(\hat{n})$ and $\hat{n} \leq N$. The payoff function of the route choice (indicated RC in Table 2.1) games is $\pi = E - c(m_j)$ for some value $E > 0$ where both the E and $c(m_j)$ can vary across actions. When E and $c(m_j)$ are identical for all actions, the game has symmetric payoff functions where switching the population between two of the available actions would not adjust the agents' payoffs. If E or $c(m_j)$ vary for even one action, then the payoff functions of the game are asymmetric. The payoff functions combine with the number of agents is used to determine the MSNE and PSNE predictions for each game.

[3] I will use iteration and run interchangeably throughout the chapter to denote an algorithm has performed all repetitions of a single play of the game.

	Experimental Games					
	D & H	C & P	Selten	Rapoport 1	Rapoport 2	Erev
Subjects	6	9	18	18	18	4
Rep	100	100	200	40 (x2)	80 (x2)	50 (x40)
Obs	3	6	6	6	6	6-9 (x40)
Game	ME	RC	RC	RC	RC	ME
Info.	Full	Full	Full	Full	Full	Limited
Sym.	No	Yes	No	Yes/No	Yes/No	No
MSNE	(1.32,4.68)	(4.5,4.5)	(12.28,5.72)	(9,9)/None	(6,6,6)/(0,0,0,9,9)	None/Various
PSNE	(2.00,4.00)	(4.0,5.0)	(12.00,6.00)	(9,9)/(0,0,18)	(6,6,6)/(0,0,0,9,9)	Various
Results	PSNE in Last 10	MSNE Dist.	PSNE Dist.	Both/Tending Dist./to PSNE	Both/Neither Dist.	Various Various

Table 2.1: Summary of Experimental Details

The symmetric payoffs produce a MSNE where each action should be taken with equal probability. Chmura and Pitz (2006) and Rapoport et al (2009) had their subjects play games with symmetric payoff functions. The other two games presented in Rapoport et al retained the symmetric payoffs, but had asymmetric payoff functions due to the addition of a new payoff function. The aggregate distribution produced in the games with symmetric payoff functions matched the MSNE prediction, but the individual behavior of the experimental subjects does not appear to match the MSNE predictions. For instance, the frequency with which the experimental subjects adjusted their play to alternative strategies does not match the prediction produced by the MSNE. The distribution predicted by the PSNE is equivalent to the MSNE prediction in the two games with symmetric payoffs found in Rapoport et al and not in Chmura and Pitz.

All of the other games used asymmetric payoffs functions and the distribution predicted by the MSNE does not match the outcomes achieved in these games. Given the failure of the MSNE, the distribution predicted by the PSNE might better describe player behavior. While the subjects in Duffy and Hopkins learned to play the PSNE over the last 10 repetitions, only one group played the PSNE for the last half of the experiment. Seven of the games found in Erev et al had the same MSNE prediction as Duffy and Hopkins. In both cases, the MSNE distribution does not provide an accurate description of the data [4]. Furthermore, none of the 46 different groups learned to play and maintain the PSNE during the first 50 periods. This suggests the PSNE may take a significant number of repetitions to learn. The PSNE also failed to predict the play in the two games used in Rapoport et al with asymmetric payoffs, but the outcome of the second game in Rapoport 1 does appear to be approaching the PSNE and the play observed in the second game of Rapoport 2 does not match either predicted outcome. Several of the games covered here do not have a MSNE with second game of Rapoport 1 representing one example and 9 of the games investigated in Erev et al representing the other examples. 73 groups played the 9 games from Erev et

[4]Using the testing described below, I could not reject the joint hypothesis for 11.3% of the data from Erev et al with the MSNE predicting entry 22% of the time

al with only 4 of the groups learning to play the PSNE during the last half of the game and 11 groups were playing the PSNE during the final 10 periods. The only game with asymmetric payoff functions where the PSNE distribution matched the outcome was Selten et al and this result only held at the aggregate level.

Given this information, it is clear that the Nash equilibrium does not describe how the agents play even if one of the (possibly many) Nash equilibrium provides a rough match to the distribution seen in the experimental data. This allows us to make predictions about what type of behavior the algorithms should display and determine how to test the predictions. An algorithm that matches the play of experimental subjects in these games should be able to produce an outcome that matches the MSNE distribution when there are symmetric payoff functions, but the algorithm should not change actions with the same frequency predicted by the MSNE. When the payoff functions are asymmetric, the algorithm should not produce an outcome that matches the distribution predicted by the MSNE. To examine whether the algorithms learn the MSNE or just match the predicted MSNE distribution, I utilize Pearson's χ^2 Goodness of Fit Test and a Z-test is applied to examine whether the agents are switching strategies as predicted by the MSNE. This means our first hypothesis tests whether the distribution created by the algorithms matches the MSNE distribution and our second hypothesis tests whether the artificial agents adjust their actions with the same frequency predicted by the MSNE, or:

1. The algorithm produces the same distribution predicted by the MSNE:

   H_0^1: $(\sum_{t=1}^{T} O_{1,t}, \dots, \sum_{t=1}^{T} O_K) = (E(N \times T \times p_1), \dots, E(N \times T \times p_K))$
   H_A^1: $(\sum_{t=1}^{T} O_1, \dots, \sum_{t=1}^{T} O_K) \neq (E(N \times T \times p_1), \dots, E(N \times T \times p_K))$

   where $O_{\ell,t}$ is the number of players selecting action $\ell \in \{1, \dots, K\}$ at time t and $E(N \times T \times f_\ell)$ is the expected outcome produced by N players using the MSNE probability p_ℓ of playing action ℓ and T is the number of repetitions of the game.

2. The algorithm switches actions with the same frequency predicted by the MSNE:

   H_0^2: $\mu = \mu_0$

   H_A^2: $\mu \neq \mu_0$

   where μ is the average number of artificial agents changing actions per period and μ_0 is the MSNE predicted number of agents changing per period.

Testing these two hypotheses alone is not sufficient to determine if the hypothesis that the algorithms are playing the MSNE should be rejected. To properly determine this, both hypotheses are tested simultaneously. The joint hypothesis test is that both H_0^1 and H_0^2 cannot be rejected, but to conduct a proper test the error rate needs to be adjusted to account for the multiple hypotheses. I determine the correct p-value based on the outcome of a simulation with all players using the MSNE[5]. All tests are conducted using a 5% p-value on the learning algorithm data produced in the second half of the games to allow the algorithms time to learn to play. The percentage of runs where the null hypothesis cannot be rejected for each of the algorithms and games is reported.

The experimental data produced several additional outcomes an algorithm mimicking human play should be able to match. To determine if an algorithm is playing the PSNE, the choices are examined to determine if they exactly matched the predicted play[6]. The percentage of iterations using the PSNE is reported for the second half of the game. When there is no MSNE, the experimental data has a small number of groups learning to play the PSNE during the later portions of the game. Thus a small percentage of the iterations should also learn to play the PSNE when there is no MSNE. The algorithms should match the distribution of play predicated by the PSNE in the game used by Selten, but the algorithm should not learn to play the PSNE. The algorithms should not match the distribution produced by both MSNE and PSNE in the second game of Rapoport 2. Finally, a measure is constructed to compare the aggregate entry/route choice produced by the human subjects with the performance of the algorithms.

2.3 The Learning Algorithms

Impulse-Matching Learning

Chmura et al (2012) proposed the Impulse-Matching Learning (IML) model. IML creates an impulse $r_\ell^i(t)$ whenever action ℓ produces a higher expected transformed payoff than action j given either action ℓ or action j is the action the agent used in the current period for player i. Each action ℓ has an impulse sum $R_\ell^i(t)$ that is the sum of all impulses from action j to action ℓ for all previous periods. Each action ℓ is played with probability $p_\ell^i(t)$ where

$$p_\ell^i(t) = \frac{R_\ell^i(t)}{\sum_{k=1}^K R_k^i(t)}$$

with K being the total number of possible actions available to each agent. And the impulse $r_\ell(t)$ in period t can be described as:

$$r_\ell^i(t) = max[0, \hat{\pi}_\ell - \hat{\pi}_j]$$

where $\hat{\pi}_\ell^i$ is the transformed hypothetical payoff from taking action ℓ and $\hat{\pi}_j^i$ is the transformed payoff from using action j. This means an impulse will exist any time an alternate action would have provided a better hypothetical payoff. The impulse sum is then updated

[5]See Appendix A.3 for details of how the adjustments are made.

[6]In other words, the χ^2 test is not used to determine if the algorithms were producing the distribution predicted by the PSNE, but whether algorithm exactly produced the predicted play for each and every period examined without deviation.

such that $R^i_\ell(t+1) = R^i_\ell(t) + r^i_\ell(t)\ \forall \ell = 1,\ldots,K$. The initial impulse sum are set to zero, i.e. $R^i_\ell(1) = 0\ \forall \ell = 1,\ldots,K$ and the probabilities are $p^i_\ell(t) = 1/K\ \forall \ell = 1,\ldots,K$ whenever $\sum_{k=1}^K R^i_k(t) = 0$.

Self-Tuning Experience Weighted Attraction

In response to criticisms of the Experience Weighted Attraction model, Ho et al (2007) adjusted the model to reduce the total number of free parameters to a single one in the Self-Tuning Experience Weighted Attraction (ST-EWA) model. This is done by adjusting the Experience Weighted Attraction (EWA) model so that one of the parameters is a fixed value and two other parameters are transformed into functions of the model itself.

Self-Tuning EWA still relies on each agent having an attraction $Atr^i_j(t)$ to a particular strategy where j is the strategy, i is the player, and t is the current period. Once again, each attraction is associated with a probability $\gamma^i_j(t) = e^{\lambda \cdot Atr^i_j(t)} / \sum_{k=1}^J e^{\lambda \cdot Atr^i_k(t)}$ of being played. The attractions can be described as:

$$Atr^i_j(t) = \frac{\phi^i(t) \cdot N(t-1) \cdot Atr^i_j(t-1) + [\delta^i_j(t) + (1-\delta^i_j(t)) \cdot \mathcal{I}_j(\hat{s}^i(t))] \cdot \pi(s^i_j(t), s^{-i}(t))}{1 + N(t-1) \cdot \phi^i(t)}$$

where $\mathcal{I}_j(\hat{s}^i(t))$ is an indicator variable equal to one if the strategy played by the agent $\hat{s}^i$ is strategy j at period t or zero otherwise, and the hypothetical utility $\pi(s^i_j, s^{-i}(t))$ is based on the payoffs described in each experiment from playing strategy $s^i_j(t)$ given the strategies $s^{-i}(t)$ used by all other players at time t where $-i$ indicates all other players. The two additional functions are the change detector function or $\phi^i(t)$ and the attention function or $\delta^i_j(t)$. The experience weight $N(t)$ is also adjusted from the previous model.

The detector function takes the form of:

$$\phi^i(t) = 1 - 0.5 \cdot SI^i(t)$$

where $SI^i(t)$ is the surprise index for each agent and is based on the difference in historical and recent behavior displayed by the other agents. $SI^i(t) = \sum_{j=1}^{\ell}(h^i_j(t) - r^i_j(t))^2$ with $h^i_j(t) = (\sum_{\tau=1}^t \mathcal{I}_j(s^{-i}_j(\tau)))/t$ is the cumulative history of the other $-i$ agents' choices to play strategy j of ℓ possible strategies and $\mathcal{I}_j$ is an indicator function taking a value of 1 if strategy j is used and 0 otherwise, $r^i_j(t) = \mathcal{I}_j(s^{-i}_j(t))$ is the recent history of agents assuming a single period memory.

The attention function is:

$$\delta^i_j(t) = \begin{cases} \frac{1}{\ell} & \text{if } \pi^i(s^i_j, s^{-i}(t)) \geq \pi^i(t) \\ 0 & \text{Otherwise} \end{cases}$$

where ℓ is the total number of strategies available to the agent and $\pi^i(t)$ is the payoff agent i received in period t. Ho et al suggested $1/\ell$ should only be used for games with mixed strategies and games with pure strategy outcomes should use 1. This means that there is a single free parameter λ to assign. I am looking for an algorithm that can function across multiple games and Chmura et al acknowledged this algorithm had the best fit for the individual level data. The same $\lambda = 0.2775$ and $1/\ell = 0.5$ found in their paper[7] are used here.

Individual Evolutionary Learning Models

Individual Evolutionary Learning is a behavioral learning model intended to address the problems associated with learning algorithms such as the Reinforcement Learning Model of Erev and Roth (1995) or the Experience Weighted Attraction Model of Camerer and Ho that performed poorly in continuous spaces, needed adjustment to the parameter space to fit the data of each experiment, and failed to approximate the dynamics seen in experimental settings (Arifovic and Ledyard, 2004). The algorithm takes advantage of evolutionary dynamics to describe an individual's learning process and attempts to find a universal set of parameters allowing the algorithm to mimic human behavior in a wide array of games. There have been several notable successes in using a consistent set of parameters that closely approximate the experimental outcomes of the Linear Public Goods Game (Arifovic and Ledyard, 2012) and Correlated Equilibria (Arifovic, et al 2015).

Each agent being modeled by the IEL plays a strategy from a set of remembered strategies. This set of strategies evolves over time by introducing new strategies and maintaining strategies that would have performed well in the past. The IEL model consists of five elements: an initial set of strategies, a method of experimentation to introduce new strategies, a forgone utility function to evaluate strategies, a method of retaining strategies that have theoretically higher associated payoffs, and a method of selecting a strategy from the set of remembered strategies. Each of the five elements will be described. Finally, the assumptions used by the IEL and the IEL with the alternative set of assumptions are discussed.

Before any evaluation can take place, a set of initial conditions is created for the computational agents to be able to play the game. Each agent is *Initialized* with a remembered strategy set A^i consisting for J strategies drawn at random from a distribution covering the set of all possible strategies. Each agent will then select a strategy $a_j \in A^i$ to be used in the first period with each element of a player's remembered strategy set having a $1/J$ chance of being used.

Once the selected strategy has been played, the agents begin to introduce new strategies via *Experimentation*. Each strategy has an independent and identical probability ρ of undergoing experimentation. If the strategy is selected, then the strategy a_j is replaced with strategy $\hat{a}_j$ drawn from a distribution limited to the set of possible strategies.

[7] I found using either 1 or $1/\ell$ for $\ell \neq 2$ produced worse outcomes for all of the games examined compared to $1/2$ and using $\lambda = 0.436$ resulted in outcomes with higher root mean square error values.

The agent then creates a set of hypothetical payoffs for each possible strategy based on the assumption that no other player changes their action. This hypothetical payoff is called the *Forgone Utility Function* $W(a_j|\tilde{a}^{-i})$ where $\tilde{a}^{-i}$ is the observed action taken by all other agents in the last period. The forgone utility evaluates the payoffs an agent would have received by using strategy a_j in the previous period holding all other players' actions constant.

Next each agent creates a new remembered strategy set $A^i(t+1)$ based on the forgone utilities $W(a_j)$ of the strategies in the current remembered strategy set $a_j \in A^i(t)$. This process is called *Replication* and involves the pairwise comparison of two strategies chosen at random $a_m, a_n \in A^i(t)$ and selecting the strategy that has the higher forgone utility to be used in the new remembered strategy set $A^i(t+1)$[8]. This process of random selection and competition is then repeated J times using all possible strategies in $A^i(t)$.

The final step in the algorithm is to determine which strategy should be used in the following period. The *Selection* of a strategy is accomplished by via a probabilistic selection rule where the probability of playing strategy $a_j \in A^i(t+1)$ is $\gamma^i_j = W(a_j)/\sum_{k=1}^{J} W(a_k)$[9].

Both algorithms are initialized by drawing each element of the remembered strategy set $a_j \in A^i(0)$ from the uniform distribution over $[0,1]$[10]. The two algorithms also rely on the same underlying learn process with each strategy giving the probability of playing an action if that strategy is selected. Once a strategy has been selected, the agent will play against a random variable $\psi \sim U[0,1]$ to determine what action to take.

The IEL is run using the same set of parameters that worked well in the past with parameter values ($J = 180, \rho = 0.033, \sigma = 0.10$). When experimentation takes place and new strategies are introduced, the IEL uses the truncated normal distribution $\hat{a}_j \sim N(a_j, \sigma)$ for $U \geq \hat{a}_j \geq L$ with U being the upper bound and L being the lower bound of the strategy set. The most literal interpretation of the forgone utility was implemented by letting $W(a_{j,\ell}|\psi, \tilde{a}^{-i}) = \pi(a_{j,\ell}|\psi, \tilde{a}^{-i})$ where π is the payoff an agent receives from playing action ℓ given the value of ψ and the actions of all other players $\tilde{a}_{-i}$[11].

The alternative set of assumptions or IEL Alt utilizes a different set of parameter values ($J = 7, \rho = 0.50, \sigma = 0.15$) determined by the grid search as described in Appendix A.2. The distribution used by the IEL Alt is the normal distribution $\hat{a}_j \sim (a_j, \sigma)$ with $\hat{a}_j$ equal

[8]A pariwise comparison is used here based on the past success of this method in other IEL models. If a tie occurs because two strategies have the same forgone utility, then one of the two strategies is chosen at random.

[9]Or if $\exists\, W(a_j) < 0$, then let $\epsilon^i = min[W(a_1), \ldots, W(a_J)]$ and adjust the probability such that $\gamma^i_j = \frac{W(a_j)-\epsilon^i}{\sum_{k=1}^{J}(W(a_k)-\epsilon^i)}$.

[10]Each strategy consists of a probability of playing each action $a_j = [a_{j,1}, \ldots, a_{j,K}]$ where K is the maximum number of actions available to the agent. Since each element of a_j is drawn from the uniform, a_j is normalized such that $\sum_{\ell=1}^{K} a_{j,\ell} = 1\, \forall\, j$.

[11]I also applied the IELORP* as described in Arifovic and Ledyard (2012) to the full information games with limited changes to the outcomes as this forgone utility function represents a form that had success in a space with continuous choice of actions as opposed to discrete choices of actions.

to the upper bound U if $\hat{a}_j > U$ or $\hat{a}_j$ equal to the lower bound L if $\hat{a}_j < L$ and $\hat{a}_j$ otherwise[12]. The hypothetical expected utility is used as the forgone utility by setting $W(a_j|\tilde{a}^{-i}) = \sum_{\ell=1}^{K} a_{j,\ell} \times \pi(\ell|\tilde{a}^{-i})$ where $a_{j,\ell}$ is the probability of strategy a_j playing action ℓ, $\pi(\ell|\tilde{a}^{-i})$ is the payoff from playing action ℓ last period, and $\tilde{a}^{-i}$ is the actions taken by all other agents.

Summary of Learning Algorithms

To provide an easier comparison of the algorithms being used, Table 2.2 summarizes the differences and similarities between the 4 algorithms. An important point to acknowledge is how the algorithms deal with the concept of a strategy. For the IML and ST-EWA each strategy is associated with using a particular action while the IEL and IEL Alt treat each strategy as a set of probabilities associated with each action. The agents are initialized so that each action is equally likely to be selected for play by IML and ST-EWA, the IEL and IEL Alt are initialized so that each action is equally likely in expectation. Once a strategy has been played, each algorithm utilizes slightly different methods to update. IML updates by increasing the summed impulse of an action by the difference in the actual payoff and hypothetical payoff when the hypothetical payoff should have been used. ST-EWA updates by discounting attraction of all actions from the previous period, and then increases the attraction of the action used and a discounted increase in the attraction of any action that would have provided a higher hypothetical payoff than the action taken. Both the IEL and IEL Alt follow a similar two step process for updating strategies. The first step has each strategy adjusted with probability ρ and using a weight of θ to determine the size of the adjustment. The second step uses the updated strategies to calculate the payoffs from the forgone utility function and then retaining strategies with higher payoffs from pairwise competition repeated for the same number of strategies the algorithm has. The final point of comparison is the method to select the strategy to be used in each period after the first. IML selects an action with a probability of the summed impulses for that action relative to the total sum of impulses over all previous periods. The ST-EWA selects an action using a Logit function to construct the probability of playing that action based on the exponential value of λ times the attraction to that action relative to the sum of all exponential values

[12] The expected mean and variance produced by this distribution can be found in Appendix A.1.

Table 2.2: Summary of Algorithm Behavior

Algorithm Name	Number of Strategies	Type of Strategies	Initial Setting	Updating Method	Selection Method
IML	Number of Actions	Probability of Playing an Action	Each Action is Equally Likely	Impulse from Transformed Difference of Action Played and the Best Action Not Played	Probability from Impulses: $p_\ell^i(t) = \frac{\sum_{m=\ell}\sum_{\tau=1}^{t-1} r_m^i(\tau)}{\sum_{\tau=1}^{t-1} r_m^i(\tau)}$
ST-EWA	Number of Actions	Probability of Playing an Action	Each Action is Equally Likely	1) All Strategies Discounted; 2) Increase Strategy Played and Strategies with a Higher Payoffs	Logit Probability from Attractions: $\gamma_j^i(t) = \frac{e^{\lambda \cdot Atr_j^i(t)}}{\sum_{k=1}^J e^{\lambda \cdot Atr_k^i(t)}}$
IEL	180	Probabilities of Playing Each Action	Each Action Equally Likely in Expectation	1) ρ Probability of Experimentation on Each Strategy; 2) Pairwise Competition (Replication)	1) Probability: $\gamma_j^i(t) = \frac{W^i(a_j(t))}{\sum_{k=1}^J W^i(a_k(t))}$ 2) Probability from Selected Strategy $a_j^i(t) = \{p_{j,1}^i, \dots, p_{j,K}^i\} \,\forall\, K$ actions
IEL Alt	7	Probabilities of Playing Each Action	Each Action Equally Likely in Expectation	1) ρ Probability of Experimentation on Each Strategy; 2) Pair Wise Competition (Replication)	1) Probability: $\gamma_j^i(t) = \frac{W^i(a_j(t))}{\sum_{k=1}^J W^i(a_k(t))}$ 2) Probability from Selected Strategy $a_j^i(t) = \{p_{j,1}^i, \dots, p_{j,K}^i\} \,\forall\, K$ actions

of λ times the attraction to each action. Both the IEL and IEL Alt utilize the same two step process where each strategy is selected with a probability based on the payoff from the forgone utility function relative to the sum of all forgone utility functions. Once a strategy is selected, both the IEL and IEL Alt use the probability of playing each action to determine which action is taken.

2.4 Results from Applying the Learning Algorithms

In this section, details of each specific game are provided, the adjustments made to each algorithm, and the performance of the algorithms is discussed. The section is divided into six parts with the first five being dedicated to the games used in each paper. The final section will give a comparison of the performance of the algorithms across papers.

When discussing specific papers, a description of the payoff functions of each game the algorithms play as well as any tools used to analyze the experimental data is provided. The algorithms were limited to using the same set of actions available to the participants of the experiment in an attempt to match the experimental setting as closely as possible. The adjustments needed to be made to the algorithms are described next, such as the transformation of payoffs used by the IML. The analysis of the algorithm's performance finishes the discussion of each paper. The performance is analyzed using the root mean squared error (RMSE) $\sqrt{\sum_{r=1}^{1000}(o^e - o^r)^2}$, where o^e is the reported outcome from the experiment and o^r is the outcome from a single run or iteration of the algorithm[13]. All of the algorithms are run on MatLab R2010a © for 1000 iterations.

The final section discusses the relative performance of each of the algorithms. This discussion targets the performance of each algorithm relative to the experimental outcomes described in Section 2. The results of the algorithm relative to the outcomes observed in the experimental data is reported[14]. The measure used to compare how well each algorithm captures the aggregate choices in each paper and across all games is described with the results reported.

2.4.1 The Market Entry Game by Duffy and Hopkins

Details of the Game

The Market Entry Game consisted of N players making a choice over *Entering* a market or *Staying Out* of the market. In the basic version of this game, *Stay Out* represented a safe choice where the payoff was independent of the number of other agents who have made the same choice. The payoff from playing *Enter* was subject to congestion with lower payoffs

[13] Further analysis and discussion for each paper is provided in Appendix A.4. For example, the normalized mean squared deviation scores for the market entry games used by Erev et al. (2010a) is reported in Appendix A.4 to allow for comparisons with the performance of other algorithms examined there.

[14] These results are based on the MSNE hypotheses testing and percentage of runs playing the PSNE

than the safe option if too many other agents also played *Enter*. Duffy and Hopkins had six subjects participate in a market entry game where the subjects were faced with the following payoff function:

$$\pi(s^i) = \begin{cases} 8 & \text{if } \textit{Stay Out} \\ 12.2 - 2 \cdot m & \text{if } \textit{Enter} \end{cases}$$

where $m =$ the number of agents choosing *Enter*. The same six subjects played the game repeatedly for 100 periods. Although there were several different informational treatments of the game, the algorithms only play the full information version of the game. The subjects were only allowed to make an indivisible choice of *Enter* or *Stay Out* in the experiment and the computational agents are limited to the same choices.

Adjustments to the Algorithms

Only the IML needs additional changes to the algorithm to account for the transformation of the game. Since the maximin outcome is v or 8, the transformed payoff function is $\hat{\pi} = 10.1 - m$ whenever $\pi > 8$ and $\hat{\pi} = 12.2 - 2 \cdot m$ otherwise.

Basic Algorithmic Results

Duffy and Hopkins (2005) focused on the experimental subjects' learning behavior. Because of their interest in learning, the paper reported data on the attendance behavior and payoffs of the agents over time. The human subject data is compared with the outcomes produced via the artificial agents using the root mean squared error as reported in Table 2.3. The IML under enters relative to the predicted PSNE outcome with the little deviation in the behavior over time. The ST-EWA over enters relative to the PSNE with the artificial agents failing to adjust their play over time. The IEL produces average outcomes near the PSNE within the first 10 periods and then fluctuates over time without ever learning to play the PSNE. The IEL Alt on the other hand over enters relative to the PSNE and over time learns to play the PSNE. The root mean square error shows how the ST-EWA maintaining the same strategy over time produces larger RMSE values over time, while the IEL Alt was able to learn and produce smaller RMSE as the artificial agents' behavior more closely resembles the experimental data. As expected, the IEL Alt has the best performance because the IEL Alt uses parameter values based on individual level data from Duffy and Hopkins.

2.4.2 The Minority Game by Chmura and Pitz

Details of the Game

In the Minority Game by Chalet and Zhang (1997), an odd number of agents N decided between two actions $\{A, B\}$ where the payoff for each agent selecting the same action as

Data	Entry			Payoff		
	All Periods	Last Half	Last 10	All Periods	Last Half	Last 10
IML	0.296	0.218	0.315	0.183	0.205	0.353
ST-EWA	0.127	0.183	0.365	0.211	0.334	0.513
IEL	0.212	0.180	0.284	0.097	0.139	0.296
IEL Alt	0.168	0.084	0.068	0.211	0.148	0.076

Table 2.3: RMSE of Algorithms compared to Duffy and Hopkins (2005)

the majority was worse than selecting the same action as the minority. Chmura and Pitz (2006) conducted an experiment using the minority game, but the game was described to the experimental subjects as a route choice game where the least crowded road provided a better payoff. The payoff function was $\pi^i(s^i = j) = 1 - round(m_j/N)$ where m_j was the number of people who played strategy $j = A$ or B. The experimental setting involved 9 subjects playing the minority game for 100 rounds. The algorithms are provided the same information as human subjects who played the full information treatment.

In addition to reporting the average outcome or route choice, the paper covered several measures of agent behavior. Chmura and Pitz (2004) described the Yule coefficient Q and utilized the resulting values to provide insight into the behavior exhibited by individuals during the game. The choices were ordered along two dimensions (i) whether the subject changed their previously played strategy and (ii) whether the subject received a payoff of H or L in the previous period as shown in Table 2.4. This grouping was then used to calculate the Yule coefficient Q as:

$$Q = \frac{q_+^s \cdot q_-^\Delta - q_-^s \cdot q_+^\Delta}{q_+^s \cdot q_-^\Delta + q_-^s \cdot q_+^\Delta}$$

The Yule coefficient Q was calculated for each individual in a session. Additionally, the authors utilized the Spearman rank correlations between the cumulative payoffs and number of road changes to further analyze the behavior of the experimental subjects.

Table 2.4: Yule coefficient Q grouping in Chmura and Pitz (2004)

	$s^i(t-1) = s^i(t)$	$s^i(t-1) \neq s^i(t)$
$\pi(s^i(t-1)\vert s^{-i}(t-1)) = H$	q_+^s	q_+^Δ
$\pi(s^i(t-1)\vert s^{-i}(t-1)) = L$	q_-^s	q_-^Δ

Adjustments to the Algorithms

Implementation of the models requires no adjustments. The IML aspiration level is zero since the available actions provide the same maximin solution and thus no transformation to the payoff function occurs.

Basic Algorithmic Results

Their paper focused on human subjects played and how each subject's road changing behavior influenced their cumulative payoffs. Table 2.5 shows that all of the algorithms are able to match the basic route choice behavior. The ST-EWA and IML produce road changing behavior closest to the experimental data. The IEL and IEL Alt fail to change roads as frequently as the experiments while the IML and ST-EWA both change roads slightly more often than the data. The Yule coefficient Q was designed to give a measure of the type of behavior exhibited by the experimental subjects and the IEL and IEL Alt gives a closer approximation than the IML and ST-EWA. The experimental data showed a statistically significant negative relationship between cumulative payoff and switching roads given by the Spearman, I find the same for both IEL and IEL Alt resulting in a smaller RMSE than the IML and ST-EWA where there is no relationship between the cumulative payoffs and number of road changes.

Data	Route A RMSE	Road Δ RMSE	Yule Coef. Q RMSE	Spearman Rank Corr. RMSE
IML	0.125	0.306	0.500	0.608
ST-EWA	0.142	0.272	0.553	0.637
IEL	0.135	0.716	0.257	0.278
IEL Alt	0.112	0.572	0.218	0.299

Table 2.5: Root Mean Squared Errors of Algorithms vs. Chmura and Pitz (2006)

2.4.3 The Route Choice Game by Selten et al

Details of the Game

Selten et al (2007) used similar measures to examine behavior as used by Chmura and Pitz. The experiment involved 18 subjects playing the game for 200 rounds. The subjects were asked to make a choice between the main road M or the side road S. The payoff for each agent was:

$$\pi(s^i) = \begin{cases} 40 - (6 + 2 * m_M) & \text{if } s^i = M \\ 40 - (12 + 3 * m_S) & \text{if } s^i = S \end{cases}$$

where s^i was the strategy of player i, m_M was the total number of agents choosing M, and m_S was the total number of agents choosing S. The Yule coefficient Q was calculated using the classification shown in Table 2.6. The full information treatment is used for comparison as this is the same information provided to the computational agents.

Table 2.6: Yule coefficient Q grouping in Selten et al (2007)

	$s^i(t-1) = s^i(t)$	$s^i(t-1) \neq s^i(t)$
$\pi(s^i(t-1)\|s^{-i}(t-1)) \geq 10$	q^s_+	q^Δ_+
$\pi(s^i(t-1)\|s^{-i}(t-1)) < 10$	q^s_-	q^Δ_-

Adjustments to the Algorithms

Since the agents are presented with two actions that have asymmetric payoffs and can results in an outcome less than zero, the IML requires a greater level of adjustment than the previous games. The possibility of achieving negative payoffs is removed by setting $\hat{E} = E - \min[\pi(m_M = N), \pi(m_S = N)] = 66$. Using 24 as the maximin payoff and $c_j(m_j)$ as the cost function of route $j \in \{M, S\}$, the IML transformed payoffs are $\hat{\pi}(s^i = j) = 45 - c_j(m_j)/2$ if $\pi(s^i = j) > -2$ and $\hat{\pi}(s^i = j) = 66 - c_j(m_j)$ otherwise.

Basic Algorithmic Results

Table 2.7 provides the root mean squared error between the basic measures reported by Selten et al and the results of the algorithms. The IEL and IEL Alt are both able to closely approximate the route choice displayed by the experimental subjects. The IML produces a result closer to the Pareto optimum level. The IEL has the smallest RMSE for road changes and produces choices closely resembling those made by the human participants. The IML again provides the worst fit relative to the data. The IML and ST-EWA produce outcomes closer to the behavior displayed in the experiments when examining the RMSE for the Yule coefficient Q.

Data	Route S RMSE	Road Δ RMSE	Yule Coef. Q RMSE
IML	0.85	4.31	0.10
ST-EWA	0.37	3.33	0.13
IEL	0.12	1.75	0.30
IEL Alt	0.18	2.49	0.27

Table 2.7: RMSE of Algorithms relative to Selten et al (2007)

2.4.4 Route Choice and the Braess Paradox by Rapoport et al

Details of the Game

The two games used by Rapoport et al (2009) represent a more complex application of the route choice game. The first game required the subjects to make a choice over two routes with each route consisting of three nodes. The experiment then added a costless connection (i.e. a four node option) and demonstrated the Braess Paradox occurring in an experimental setting with human subjects. The 18 subjects played each part of the game repeatedly over 40 periods (i.e. 80 total periods). The payoffs for the two route game (or game $1_{RC=2}$) were:

$$\pi(s^i) = \begin{cases} 400 - (210 + 10 * m_A) & \text{if } s^i = O - A - D \\ 400 - (210 + 10 * m_B) & \text{if } s^i = O - B - D \end{cases}$$

while the three route game (or game $1_{RC=3}$) had payoffs of:

$$\pi(s^i) = \begin{cases} 400 - (210 + 10 * (m_A + m_{A-B})) & \text{if } s^i = O - A - D \\ 400 - (210 + 10 * (m_B + m_{A-B})) & \text{if } s^i = O - B - D \\ 400 - (10 * (m_A + m_{A-B}) + 10 * (m_B + m_{A-B}) & \text{if } s^i = O - A - B - D \end{cases}$$

where m_A was the number of agents choosing route $O-A-D$, m_B was the number of agents choosing route $O-B-D$, and m_{A-B} was the number of agent choosing route $O-A-B-D$. Notice that route $O-A-B-D$ simply added a costless connection between two previously unconnected nodes eliminating the fixed cost component. The Braess Paradox states adding a costless connection will in fact lead to a decrease in the overall welfare of the players (i.e. a lower average payoff) if the connection joins nodes where the costs are variable. To ensure the support for the Braess Paradox was robust and not due to the order the games were played, the experiments were repeated with the three route choice option presented first and the costless route removed.

The second experiment presented in Rapoport et al focused on 18 subjects deciding over 3 possible routes for 80 rounds, then introduced two additional costless connections and had the subjects choose over 5 possible routes for 80 rounds. Given the lack of difference in behavior from the order the games were played in the first experiment, this game was only run with the costless connections added. The three route choice payoffs (or game $2_{RC=2}$) were:

$$\pi(s^i) = \begin{cases} 196 - (60 + 12 * m_A) & \text{if } s^i = O - A - D \\ 196 - (60 + 12 * m_B) & \text{if } s^i = O - B - D \\ 196 - (60 + 12 * m_{C-E}) & \text{if } s^i = O - C - E - D \end{cases}$$

while the five route game (or game $2_{RC=5}$) had payoffs of:

$$\pi(s^i) = \begin{cases} 196 - (60 + 12 * (m_A + m_{C-A})) & \text{if } s^i = O - A - D \\ 196 - (60 + 12 * (m_B + m_{B-E})) & \text{if } s^i = O - B - D \\ 196 - (60 + 12 * m_{C-E} + 6 * (m_{C-A} + m_{B-E})) & \text{if } s^i = O - C - E - D \\ 196 - (2 + 12 * (m_{C-A} + m_A) + 6 * (m_{C-A} + m_{C-E})) & \text{if } s^i = O - C - A - D \\ 196 - (2 + 12 * (m_{B-E} + m_B) + 6 * (m_{B-E} + m_{C-E})) & \text{if } s^i = O - B - E - D \end{cases}$$

where n_A was the number of agents choosing route $O-A-D$, n_B was the number of agents choosing route $O-B-D$, m_{C-E} was the number of agents choosing route $O-C-E-D$, m_{C-A} was the number of agents choosing route $O-C-A-D$, and m_{B-E} was the number of agents choosing route $O-B-E-D$. The two costless connections were added from node C to node A and from node B to node E that were previously unconnected routes.

Adjustments to the Algorithms

This set of experiments requires the most adjustments because of the additional actions available to the artificial agents. The IEL and IEL Alt are adjusted to account for how experimentation influences the probabilistic choices made. The ST-EWA requires no changes or adjustments. The IML requires an adjustment to determine what action is influenced by the impulse and the payoff transformation. The section concludes by describing how the algorithms are adjusted to address each game consisting of two parts.

Both the IEL and the IEL Alt use strategies that represent the probability of playing an action, thus each strategy consists of up to 5 elements that sum to one. Since experimentation introduces new strategies by adjusting one probability, this creates a need for greater adjustment when more than two actions exist[15]. Let Q be a random integer between 1 and the length of the strategy(i.e. number of actions). If the rule is selected for experimentation, then the Q^{th} rule is adjusted using the distribution from Section 2.3. The sum of the elements of the strategy can now be greater than or less than one, so the strategy is rescaled by dividing each probability by the sum of the probabilities. The effects of experimentation are smaller due to the rescaling.

The IML has only been applied in games with two actions. This leads to the question of which of the actions not played should determine the size of the impulse and should the impulse influence one action or all actions. The two options are for either all actions not taken being influenced by an impulse relative to the action taken or only the best action not taken being influenced (i.e. the opportunity cost). The second option is implemented[16] for the IML creating the following influence of the impulse:

$$r_i(t) = \begin{cases} max[0, \hat{\pi}_i - \hat{\pi}_j] & \text{where } \hat{\pi}_i > \hat{\pi}_k \; \forall k \neq i, j \text{ and } j \text{ is the action used} \\ max[0, \hat{\pi}_i - \hat{\pi}_j] & \text{where } \hat{\pi}_j > \hat{\pi}_k \; \forall k \neq i, j \text{ and } i \text{ is the action used} \\ 0 & \text{otherwise} \end{cases}$$

For game $1_{R=2}$, no transformation of the payoffs is made since the actions have identical maximin payoffs and all payoffs are greater than zero. Game $1_{RC=3}$ requires a transformation of the payoff $\hat{\pi}(s^i = j|s^{-i}) = 220 - c(m_j|m_{-j})/2$ for $\pi(s^i = j|s^{-i}) > 40$ and $\hat{\pi}(s^i = j|s^{-i}) = \pi(s^i = j|s^{-i})$ otherwise with s^{-i} being the strategy used by all other

[15] The strategy in the two action case consisted of the probability of playing action one p_1 and the probability of playing action two$p_2 = 1 - p_1$, making the adjustment from experimentation a direct one-to-one change.

[16] In the case of ties, the action influenced by the impulse is randomized amongst the tied outcomes.

players, m_{-j} being a vector containing the actions taken by all other players, and an aspiration level of 40. The payoffs are transformed for game $2_{RC=3}$ to avoid the possibility of a negative payoff. Since the payoff functions all have the same minimum payoff, the transformed payoffs are $\pi(s^i \hat{=} j|s^{-i}) = 80 + \pi(s^i = j|s^{-i})$ in game $2_{RC=3}$. Game $2_{RC=5}$ requires subtracting the minimum payoff -130 to ensure non-negative payoffs. Since the previous minimum payoff of -80 is still achievable by the IML agents, this creates an aspiration payoff of 50 and transformed payoffs $\hat{\pi}(s^i = j|s^{-i}) = 253 - c(m_j|m_{-j})/2$ for $\pi(s^i = j|s^{-i}) > 50$ and $\hat{\pi}(s^i = j|s^{-i}) = \pi(s^i = j|s^{-i})$ otherwise.

Another issue for the algorithms is the question of how the choices made in the first part of the game influences the choices made in the second part of the experiment. One option is to have the information learned in the previous game retained and utilized to determine what actions are used when the costless connection is added or deleted. This option is not used because unlike the experimental results retaining the information resulted in most of the algorithms producing outcomes that differed significantly when the node is added versus when the node is deleted[17]. The other option is to assume the algorithms play as if the prior game had not occurred. While this assumption produces behavior consistent with the experimental data [18], any analysis of behavior between the two games is spurious.

Basic Algorithmic Results

Table 2.8 provides the root mean squared error of the choice of route and the number of road changes per agent. The RMSE for m_B in game $1_{RC=2}$ is not reported since it matches the value reported for route A. The IML, IEL, and IEL Alt all produce outcomes similar to the experimental data for route choice in game $1_{RC=2}$, but fail to match the road changes with the IEL Alt coming closest to matching. The ST-EWA has the highest RMSE for both items examined for game $1_{RC=2}$ because almost all the artificial agents the ST-EWA algorithm creates stop adjusting their play after the 5^{th} period of play. Since all of the agents in both the IML and the ST-EWA start playing the PSNE in game $1_{RC=3}$ by the second period of play, the RMSE for the IML and ST-EWA are similar. The IEL is slowly moving toward the PSNE, creating a large difference between how the algorithm performs and the experimental data. Only the IEL Alt has a dynamic similar to the experimental data in adjusting toward the PSNE with a small number of the iterations playing the PSNE.

[17]Only the IEL Alt performs essentially the same in both games whether the costless connection is added or deleted where the extra probability is introduced as a random variable or the extra probability is removed from the remembered strategy set. The IEL produces significantly different outcomes between the two conditions using the same process. The other two algorithms are too slow to adjust because of all of the weight already accumulated when the costless connection is added, but produced nearly identical data in game $1_{RC=2}$ regardless of adding or subtracting the costless connection.

[18]To be clear, this means the data are virtually indistinguishable in the two experiments whether the connection was added or deleted and is not claiming the artificial agents produce data identical to the experiment

Data Game $1_{RC=2}$		RMSE m_A		RMSE Road Δ
IML		0.11		8.87
ST-EWA		1.08		11.65
IEL		0.27		4.47
IEL Alt		0.22		4.28
Data Game $1_{RC=3}$	RMSE m_A	RMSE m_B	RMSE m_{A-B}	RMSE Road Δ
IML	1.55	1.34	2.89	9.39
ST-EWA	1.49	1.23	2.76	9.37
IEL	1.26	1.33	2.58	7.48
IEL Alt	0.52	0.30	0.78	2.99

Table 2.8: RMSE in the ADD version of Game $1_{RC=2}$ and Game $1_{RC=3}$

The root mean squared errors comparing the route choice and number of road changes per player are shown in Table 2.9. The IML, IEL, and IEL Alt all produce similar RMSE for routes in game $2_{RC=3}$. The IEL Alt most closely matches the road changes, but adjusts much less frequently than human subjects. The RMSE for the road changes of the IEL is similar to the IEL Alt with the exception that the IEL changes too frequently. The IML also adjusts more frequently than the human subjects. The ST-EWA once again has a large number of agents playing the same strategy with a majority playing the PSNE and others maintaining the same out of equilibrium outcomes. Game $2_{RC=5}$ represents the most complex game examined. The IEL Alt produces the lowest RMSE for almost all of the outcomes suggesting it provides the closest fit to the data. The IML and ST-EWA have similar RMSE for route choice with the ST-EWA producing almost zero average road changes per player. The RMSE for road changes is lower for the IML than the IEL Alt, although the values are similar. The IEL provides the worst fit for the route choice data in game $2_{RC=5}$, but this is based on the overly simplified forgone utility choice.[19].

2.4.5 The Market Entry Game by Erev et al

Details of the Game

Erev et al (2010a) presented experimental data of another market entry game as well as the initial portion of a competition examining the ability of algorithms to match the experimental data. Unlike the previous market entry game, this version included an element of uncertainty about the outcome of the game. Additionally, the subjects participating in the

[19]To determine how influential the choice to use the simplest possible forgone utility function, the algorithm is re-run allowing for IEL to exhibit more complex behavior. The adjustments includes breaking ties of the forgone utility by determining the best forgone utility from the previous period(s), using the hypothetical expected utility as the forgone utility, and mimicking the experimentation/forgone utility of the IEL Alt. Despite all of these adjustments the IEL Alt continues to outperform the IEL, although the IEL is no longer the algorithm with the worst fit after these adjustments.

Data Game $2_{RC=3}$	RMSE m_A	RMSE m_B	RMSE m_{C-E}	RMSE Rd. Δ
IML	0.20	0.25	0.20	18.48
ST-EWA	0.53	0.54	0.54	34.35
IEL	0.24	0.20	0.23	10.80
IEL Alt	0.15	0.14	0.15	9.41

Data Gm. $2_{RC=5}$	RMSE m_A	RMSE m_B	RMSE m_{C-E}	RMSE m_{C-A}	RMSE m_{B-E}	RMSE Rd. Δ
IML	1.07	0.90	1.11	1.55	1.47	12.43
ST-EWA	1.02	1.12	0.92	1.35	1.46	34.66
IEL	1.13	1.32	1.61	1.87	2.16	20.91
IEL Alt	0.24	0.35	0.37	0.37	0.53	13.36

Table 2.9: RMSE in Game $2_{RC=3}$ and Game $2_{RC=5}$

experiment were not provided the full information of the game. Feedback to the participants was limited to their actual payoff and the hypothetical payoff had the other action been selected.

The experiment introduced environmental uncertainty in the form of a gamble $G(t)$ that had a probability ϱ of giving a high payoff H and $1-\varrho$ of giving a low payoff L. This gamble had an effect on both those who entered the market and those who stayed out of the market. The environmental effect was lower on those who stayed out of the market since the gamble was reduced by $S > 1$. The payoffs were:

$$\pi(s^i) = \begin{cases} 10 + G(t) - K \cdot m & \text{if } Enter \\ \text{round}(G(t)/S) & \text{if } Stay\ Out \text{ and } p \leq 0.5 \\ -\text{round}(G(t)/S) & \text{if } Stay\ Out \text{ and } p > 0.5 \end{cases}$$

where K was a whole number representing the cost of entering, m was the number of agents playing *Enter*, and p was a random variable drawn from the uniform distribution between zero and one. The data consisted of 120 students playing 10 versions of this game[20] for 50 periods.

Adjustments to the Algorithms

Since the participants are only provided actual and hypothetical payoffs based on the two actions available, the algorithms are limited to the same information. While this information matches the information normally given to the ST-EWA, IEL, and IEL Alt, the IML requires full information to determine the aspiration level and properly transform the payoffs; however, calculating the maximin payoff is impossible since the agents are unaware of the game's structure and the IML is run assuming an aspiration level of zero. The payoffs

[20] See p. 120-121 of Erev et al for a listing of the 40 possible game combinations with varying $\{H, L, \varrho, K, S\}$.

are still transformed to $\hat{\pi}_i = \pi_i - \min[0, \pi_i, \pi_j]$ for actions i, j in order to avoid negative outcomes.

Basic Algorithmic Results

The results reported by Erev et al focused on the Normalized Mean Squared Deviation while this chapter focuses on data consistent with the outcomes of other games. Because all 40 games are grouped together, Table 2.10 provides the mean of the root mean squared error for the average payoff, the total entry, the total number of times an agent switched actions, and the aggregate outcome of the Spearman rank correlation between the cumulative payoff and the total number of switches made by an agent [21]. The algorithms all produce payoffs that are similar to those produced in the experiment with the IEL Alt being the furthest from the actual results. The RMSE for Enter is based on the total number of experimental participants who entered relative to the actions of the algorithm. While not shown on the table, the ST-EWA is the only algorithm with over entry at the aggregate level. All of the other algorithms produce an average entry rate below the experimental data.[22] In other words, the ST-EWA averages 5.82 more agents entering than in the experiment; conversely, the IML averages 5.81 fewer agents entering than in the experiment. The IEL Alt produces the worst fit of the entry data shown by the large RMSE. The IEL Alt produces the best (smallest) result with regard to how many changes agents are making and is the only algorithm to produce fewer changes than seen in the data. All of the other algorithms adjust their actions more frequently. The final comparison is the Spearman rank correlation with all of the algorithms producing relatively similar outcomes for the RMSE of the Spearman rank correlation.

Data	Average Payoff RMSE	Enter RMSE	RMSE of Strategy Changes	Spearman Rank Corr. RMSE
IML	1.98	5.81	9.22	0.18
ST-EWA	2.19	5.82	10.10	0.23
IEL	2.16	7.03	6.32	0.15
IEL Alt	2.24	8.17	5.00	0.14

Table 2.10: RMSE of Algorithms relative to Erev et al (2010b)

2.4.6 Comparisons of the Algorithms across Games

The analysis thus far has been limited to the ability of the algorithms to replicate the results of a single game. While this is beneficial for describing algorithm behavior through multiple measures, it does not provide information about how the algorithms perform across

[21] The Spearman rank correlation is -0.27 and significantly different from zero with $p < 0.01$.

[22] The ST-EWA is the only algorithm producing this result in the market entry game just as it produced over entry when playing the game used by Duffy and Hopkins.

games. To compare the overall performance of the learning algorithms across games, the algorithms are examined to determine their ability to match the outcomes noted in Section 2 and the aggregate route/entry choice. The prediction in Section 2 is the outcome will match the distribution predicted by the MSNE when the payoff functions are symmetric, but the algorithm should produce an outcome statistically different from the MSNE due to the frequency of agent switching. Whenever the payoff functions are asymmetric, the outcome produced by the algorithm should be statistically different from the distribution predicted by the MSNE. Since all of the simulated behavior is based on learning algorithms, only the data from the last half of each game is examined to allow the algorithms time to learn. Finally, the measure used to compare the ability of the algorithms to capture the aggregate route/entry choice is defined and reported.

Summary of Outcomes Relative to the Observed Experimental Data

In this section, the ability of the algorithms to match patterns observed across games is examined[23]. In the symmetric games, the experimental subjects produced a distribution consistent with playing the MSNE, but the behavior at the individual level was consistent with using the MSNE. The overall outcomes produced in the experiments when payoffs were asymmetric are more complex with the outcomes generally differing from any of the Nash Equilibria. Table 2.11 details the percentage of runs matching the behavior observed in the congestion games when payoffs are symmetric or asymmetric.

The games with symmetric payoffs examined here all produced a distribution consistent with the MSNE, but the relative frequency of strategy changes by individuals was not consistent with the MSNE. Table 2.11 reports how the algorithms perform as a percentage of iterations that are statistically consistent with the pervious sentence, i.e. the percentage should match the “Expected” line. The IML is generally very successful at learning to play the MSNE in these games, but the algorithm does learn to play the PSNE approximately 20% of the time in the Rapoport et al 1_{RC_2} (R et al $1_{RC=2}$ in the Table) experiment. The ST-EWA only learns to play the MSNE in the C & P game, whereas the algorithm fails in the two Rapoport games because it learns to play the PSNE. The IEL produces outcomes consistent with the experimental data over 95% of the time in all of the games where payoffs are symmetric. The IEL Alt performs relatively well across all of the games with the biggest failure to match the data in the C & P game.

The games with asymmetric payoffs generally produced more ambiguous outcomes, but generally do not produce outcomes consistent with the MSNE. Table 2.11 gives the percentage necessary to be consistent with the outcomes reported in the papers as well as the percentage of iterations of the algorithm that are consistent with the behavior observed in the experiments. D & H reported that 1/3 of the outcomes were consistent with players

[23] Appendix A.4 provides a detailed analysis of the outcomes using the hypothesis testing discussed in Section 2. Additionally, the analysis of Erev et al can be found in this section.

	Experimental Games: Symmetric Payoffs		
Algorithm	C & P	R et al $1_{RC=2}$	R et al $2_{RC=3}$
Expected	100.0%	100.0%	100.0%
IML	3.0%	2.1%	5.6%
ST-EWA	4.8%	68.5%	45.9%
IEL	96.8%	97.5%	99.7%
IEL Alt	84.2%	100.0%	99.8%

	Experimental Games: Asymmetric Payoffs			
Algorithm	D & H	R et al $1_{RC=3}$	R et al $2_{RC=5}$	S et al
Expected	33.3%	100.0%	100.0%	33.3%
IML	24.5%	0.0%	100.0%	0.0%
ST-EWA	0.0%	1.0%	100.0%	0.0%
IEL	0.0%	100.0%	100.0%	67.4%
IEL Alt	15.1%	99.8%	100.0%	20.0%

Table 2.11: Performance Relative to Experimental Games

learning to play the PSNE during the second half of the game. Both the IML and IEL Alt have a positive percentage of iterations learn to play the PSNE while none of the iterations of the ST-EWA and IEL learn to play the PSNE. In Rapoport et al $1_{RC=3}$ the experimental subjects failed to learn to play the PSNE[24], while the IML and ST-EWA generally learn to play the PSNE with the IEL and IEL Alt generally failing to learn to play the PSNE. Rapoport et al $2_{RC=5}$ failed to have an outcome consistent with either the MSNE or the PSNE and all of the algorithms matched this result. Selten et al (S et al in Table 2.11) found that 1/3 of the runs were consistent with the MSNE distribution. Only the IEL and IEL Alt produced outcomes that were consistent with the MSNE distribution while the IML and ST-EWA did not produce any outcomes consistent with this result.

Aggregate Choice across Games

All of the games had the human subjects make an observable choice over a limited number of actions. This set of choices made over time was the most important variable in determining the equilibrium outcome produced in the experiment. Due to the differing number of agents and actions available to the agents, a direct comparison of behavior across games is fruitless. Instead a measure is implemented to allow comparisons of the performance of the algorithms across the games by accounting for the different number of agents, actions, and repetitions. The measure for the aggregate choice is similar to the mean quadratic distance found in Chmura et al; however, their measure relied on the symmetry of the 2 player 2×2 games. The same symmetry does not exist in the games analyzed here because the agents are not classified as row or column players and there are more than 2 actions for 3 of the games. The measure used is:

[24]The MSNE does not exist in this game.

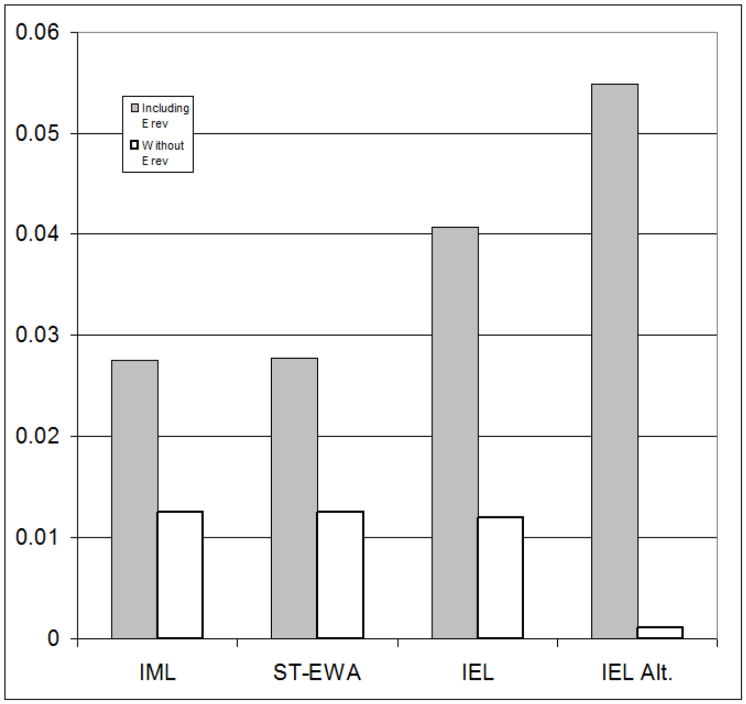


Figure 2.1: Average Sum of Squared Distance Over All Games with and without Erev Included

$$\text{Average Sum of Squared Distance} = \sum_{g=1}^{G}(\sum_{r=1}^{R}(\sum_{\ell=1}^{L}(\frac{o^{e}_{\ell,g}}{N} - \frac{o^{r}_{\ell,g}}{N})^2))/(R*G)$$

where G is the number of games played by the algorithm, R is the number of runs of the algorithm, L is the number of actions available to the players, N is the number of agents in the game, $o^{e}_{\ell,g}$ is the average outcome of action ℓ and game g in the experiment, and $o^{r}_{\ell,g}$ is the average outcome of action ℓ and game g produced by iteration r of the algorithm. The measure has a minimum of zero implying the algorithm is able to perfectly match the data and a maximum of two implying the algorithm produces an outcome diverging from the data.

Algorithm	Experimental Games				
	D & H	C & P	Selten	Rapoport	Erev
IML	0.0049	0.0004	0.0045	0.0170	0.0314
ST-EWA	0.0009	0.0005	0.0008	0.0185	0.0306
IEL	0.0025	0.0005	0.0001	0.0174	0.0472
IEL Alt	0.0016	0.0003	0.0002	0.0013	0.0668

Table 2.12: Average Sum of Squared Distance

Table 2.12 shows the score achieved by each algorithm in each of the five papers. The IEL Alt produces outcomes that generally match the entry rates obtained in the experiments when the agents have full information. Figure 2.1 shows the overall average sum of squared distance including the data from all of the games and when the outcome when the Erev games are excluded. This figure clearly shows that the IEL Alt on average is able to approximate the aggregate performance of human subjects better than the other learning algorithms in the games where the agents had access to full information. The previous section details how the IEL Alt produces outcomes consistent with the predicted outcomes from the experimental data for the full information games; however, the IEL Alt fails to perform as well as the other algorithms in the 40 games used by Erev et al. On average, the IEL performs as well as the other three algorithms as seen in figure 2.1 when not including the Erev games, but does slightly worse than the ST-EWA and IML when the Erev data is included in the average. Overall the IML and ST-EWA produce similar average outcomes as shown in Figure 2.1. The ST-EWA produces the outcomes closest to the aggregate choices of the human subjects in the two market entry games and is the only algorithm where there is over entry at the aggregate level just like the experimental data.

2.5 Concluding Remarks

The purpose of the study is to determine if any of the algorithms examined are able to mimic the regularities produced by human subjects across a range of games with congestion. This chapter reports the RMSE for each algorithm relative to the experimental data from each paper, reports the ability of each algorithm to produce a play consistent with the MSNE and PSNE predictions, and reports a measure to score the performance of each algorithm. Each of these items is used to determine if the algorithms could match the regularities produced by human subjects. The Individual Evolutionary Learning Alternative is able to produce outcomes statistically indistinguishable from the MSNE distribution when the game has a symmetric payoff function, is able to generate outcomes statistically different from the MSNE distribution when the game has asymmetric payoff functions, and on average matches the aggregate experimental data. While the IEL is designed to deal with complex strategy spaces, it performs well in a discrete choice environment as long as the number of actions is limited. A more sophisticated forgone utility function allows the IEL to improve the overall performance in games where the discrete actions are more numerous, but the IEL with a simplified forgone utility function is still able to produce outcomes comparable with the other learning algorithms. The IML requires extra effort to determine the aspiration level and transform the game; however, the lack of parameter choice does provide a clean and simple first approximation of behavior. I also show methods to transform the payoffs when payoffs are negative and/or there exists more than two possible actions. Unfortunately, the IML is prone to finding the PSNE in the first period and then maintaining this outcome indefinitely, a behavior never displayed by the experimental subjects. Finally the ST-EWA provides some of the least realistic outcomes with the artificial agents often finding and maintaining the same action overtime. The results of this chapter provides a useful examination into the behavior of each of the algorithms expanding our understanding of the conditions when the algorithm is useful as a first approximation of the outcome prior to running an experiment.

Chapter 3

Learning Correlated Equilibria: An Evolutionary Approach

(with Jasmina Arifovic and John Duffy)

3.1 Introduction

Correlated equilibrium, as first proposed by Aumann (1974, 1987), represents an important generalization of the Nash equilibrium concept. In a correlated equilibrium, players' beliefs are correlated with external signals that follow known probability distributions. Thus, in a correlated equilibrium, beliefs are not probabilistically independent as they must be in the mixed strategy interpretation of Nash equilibrium, but are instead correlated. A mutual best response to such correlated beliefs comprises a correlated equilibrium. Since extraneous signal realizations can matter in equilibrium, correlated equilibrium is considered a more general version of equilibrium behavior in non-cooperative games than is Nash equilibrium where such extraneous signals can play no role. Indeed, as Myerson has reportedly observed:

> "If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium."

The presence of external signals, however, presents a further puzzle: how is it that agents know the distribution of such external signals and use that knowledge to correlate their beliefs with signal realizations in such a way as to play a correlated equilibrium of the game? In this chapter we take a first step toward addressing this question. We develop an agent-based, evolutionary learning model with the aim of demonstrating (and in the process understanding) how boundedly rational adaptive agents with no prior knowledge of the distribution of external signals could learn to use those signals as a correlation device for their beliefs and how signal realizations map into action choices in a non-cooperative game. To provide some external validity for our simulation exercises, we compare our results with those of experimental studies involving paid human subjects playing the same coordination games used in our agent-based simulations.

An understanding of how agents might learn to play a correlated equilibrium also has important policy implications if one takes the view that institutions or policy responses depend upon the use of correlation devices (Gintis, 2009). According to this view, the efficiency advantages of all types of institutions or policy responses arise from the decisions of agents to rationally follow correlated signals. For example, a traffic signal allows for a much more efficient flow of traffic than if drivers were to navigate road intersections without such correlating devices. Alternatively, one may think of the correlating device as shocks to the macroeconomy, to which fiscal and monetary authorities have to coordinate their policy responses.

3.2 Related literature

The literature on correlated equilibrium begins with Aumann (1974, 1987) who showed that correlated equilibrium is the natural equilibrium concept under Bayesian rationality, and is less demanding than Nash equilibrium in that it does not require common knowledge of rationality. Correlated equilibrium has also been shown to be the outcome of non-cooperative games with pre-play communication (e.g., Forges, 1990; Ben-Porath,1998; and Moreno and Wooders, 1998). There are also some theory papers providing conditions (or learning rules) under which correlated equilibria can be learned in the limit of an infinitely repeated game provided that players adhere to certain regret-minimizing strategies as first demonstrated by Foster and Vohra (1997) and elaborated upon by Fudenberg and Levine (1999), Hart and Mas-Colell (2000) and Lehrer (2003) among others. These approaches all presume that agents can perfectly recall the entire past history of play of the game by all potential opponents. By contrast, our evolutionary learning approach does not require such extreme informational requirements. Further, we are interested in validating our learning results by matching them with experimental findings exploring play of correlated equilibrium strategies in *finitely* repeated games. Indeed, we find that our individual evolutionary learning model yields a good fit to the outcomes of two experimental studies on correlated equilibrium.

There are many studies using agent-based evolutionary models that seek to understand how agents learn over time to coordinate on a Nash equilibrium or to select from among several Nash equilibria. However, we are not aware of any agent-based models exploring the learning of correlated equilibria, even though as noted earlier, it is a more general solution concept.

Experimental studies of the correlated equilibrium concept have been conducted by Cason and Sharma (2007), Duffy and Feltovich (2010), Bone et al. (2013) and Duffy et al. (2016). Here we compare the results of simulations using our individual evolutionary learning algorithm to experimental findings for play of the Battle of the Sexes game by Duffy et al. (2016) and play of a Chicken game by Duffy and Feltovich (2010).

3.3 Correlated Equilibria in Two Games

Our analysis makes use of two, two-player non-cooperative games that are often studied in the literature on correlated equilibrium, the Battle of the Sexes game shown in Table 3.1 and the Chicken game shown in Table 3.2.

		Player 2	
		C	D
Player	C	9,3	0,0
1	D	0,0	3,9

Table 3.1: Battle of the Sexes

		Player 2	
		C	D
Player	C	7,7	3,9
1	D	9,3	0,0

Table 3.2: Chicken

In both games there are two pure strategy equilibria and one symmetric mixed strategy equilibrium. In particular, let $\{p_1,\ p_2\}$ denote the probabilities with which players 1 and 2 play strategy C, so $1-p_i$ is the probability that i plays strategy D. Let the payoff from each strategy be given by the pair $\{\pi_1, \pi_2\}$. Then, the Nash equilibria (NE) can be summarized as follows:

Game	Pure NE1 / Payoffs	Pure NE2 / Payoffs	Mixed NE /Expected Payoffs
BoS	$\{1,1\}$ / $\{9,3\}$	$\{0,0\}$ / $\{3,9\}$	$\{.75,.25\}$ / $\{2.25,2.25\}$
Chicken	$\{1,0\}$ / $\{9,3\}$	$\{0,1\}$ / $\{3,9\}$	$\{.60,.60\}$ / $\{5.40,5.40\}$

In the absence of external signals, experimental subjects who are randomly and anonymously paired to play these games generally coordinate on the mixed strategy Nash equilibrium (Cason and Sharma (2007), Duffy and Feltovich (2010)).

However, suppose there is an external randomization device or, as in Myerson (1991), there is a third party mediator who makes recommendations to both players as to how they should play the game in each period. One simple device is a coin flip: If heads, the recommendation is that Player 1 plays C, and Player 2 plays C, with the opposite recommendation if the coin flip is tails, as summarized in this table:

Recommended Play		Player 2 C	Player 2 D
Player 1	C	1/2	0
	D	0	1/2

Players always know what the mediator recommends for them to play and may also know what the mediator recommends for their opponent to play but the latter information is not required; it suffices that the recommendations are correlated with one another according to the stochastic external recommendation process. Suppose that players follow the action recommended for them by the mediator. In this "turn-taking" equilibrium players do better than in the symmetric mixed-strategy equilibrium, earning an expected payoff of 6 (rather than 2.25). The reason is straightforward; following recommended actions (like following traffic signals) allows players to avoid "miss-coordination" on 0-payoff outcomes, which is unavoidable when subjects play according to the symmetric mixed strategy Nash equilibrium.

In the game of Chicken, a turn-taking equilibrium of the same sort also exists. But in Chicken, by contrast with Battle of the Sexes, players can do even better using third party recommendations. For example, suppose the distribution of recommended strategy profiles is

Recommended Play		Player 2 C	Player 2 D
Player 1	C	1/3	1/3
	D	1/3	0

and players only see their part of the recommended strategy profile and not the other players' part. In that case, it is an equilibrium best response to play according to the recommended actions: a player's expected payoff from doing so is $6\frac{1}{3} > 6$, the payoff earned in the turn-taking correlated equilibrium, which continues to dominate the payoff in the symmetric mixed strategy equilibrium (5.40). This example illustrates the possibility of generating correlated equilibria with payoffs that lie outside the convex hull of payoffs possible from randomizations over the set of pure strategy Nash equilibria.

More generally, let the set of possible joint actions for players $i = 1, 2, \cdots N$ in game Γ be denoted by $\mathcal{A}$, and let $p = \Delta(\mathcal{A})$ denote some probability distribution over $\mathcal{A}$. Further, let p^a denote the probability of joint action profile $a = (a^i, a^{-i})$ where $\sum_{a \in \mathcal{A}} p^a = 1$. The probability distribution p is a *correlated equilibrium* of the game Γ having payoffs $\pi(a)$ if for all players $i \in N$ and all actions a^i, $\hat{a}^i \in \mathcal{A}$ we have that:

$$\sum_{a^{-i} \in \mathcal{A}} \pi(a^i, a^{-i}) p^{(a^i, a^{-i})} \geq \sum_{a^{-i} \in \mathcal{A}} \pi(\hat{a}^i, a^{-i}) p^{(a^i, a^{-i})}.$$

In words, given some known probability distribution, p, over action profiles, a deviation from the recommended component of player action profile is unprofitable for any player

i. Notice two features of this equilibrium concept. First, there is no requirement that the probability distributions are independent across players as in the standard Nash equilibrium definition. Second, there can be many different correlated equilibria of the game Γ, namely any probability distribution over recommended strategy profiles that satisfies the condition given above. It is this sense that correlated equilibrium can be regarded as a generalization of the Nash equilibrium concept.

The recommendations used in this study will follow those used in Duffy et al. (2016) and Duffy and Feltovich (2010) to achieve a correlated equilibrium in the Battle of the Sexes and the Chicken games, respectively. These correlated equilibria can be identified by a randomization device, $\mathcal{D}=(\Omega, \{H_i\}, \pi)$ where Ω is a finite state space corresponding to the outcomes of the randomization device, H_i is Player i's information partition on Ω, and π is a probability measure on Ω. As in the human subject experiments, we consider a class of devices $\mathcal{D}$ with the following profile of state space and partitions:

1. A state space with four elements, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
2. Player 1's information partition is $H_1 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\}$ and player 2's is $H_2 = \{\{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}\}$.

For example, in the BoS game, $\{\omega_1\} = (C, C)$, $\{\omega_2\} = (C, D)$, $\{\omega_3\} = (D, C)$, and $\{\omega_4\} = (D, D)$. Using the turn taking example discussed above for Table 3.2 would have the randomization device be $r = Pr(\omega_1) = Pr(\omega_4) = \frac{1}{2}$, i.e., the randomization device is equally likely to recommend for player 1 to play C and player 2 to play C as it is to recommend for player 1 to play D and for player 2 to play D. The particular randomization devices used in BoS and Chicken will be described below.

3.4 Individual Evolutionary Learning Model and Results

In this section we describe an evolutionary approach to learning correlated equilibria in the Battle of Sexes and Chicken games described in the previous section. We also report on simulations using our model in each game. Our objective is to model a process that might allow artificial agents with bounded memory and no prior knowledge of the distribution of external recommendations or of the mapping from recommendations to action choices to eventually learn to play a correlated equilibrium in a finitely repeated game.

Specifically, we adopt and modify the Individual Evolutionary Learning (IEL) model of Arifovic and Ledyard (2004) which views the evolutionary learning process as operating at the individual player level. The IEL model has been used in many different environments to successfully explain behavior observed in experiments with human subjects.[1] In this section we show how IEL can be extended to incorporate external correlation devices.

In the basic IEL model, each player i has a finite set of remembered strategies, S^i. The payoff function for player i, $\pi^i(s^i, s^{-i}) \rightarrow \Re$ is used to determine the results of the stage game when strategy profile $s \in S$ was chosen by all players. The set of remembered strategies, S, evolves over time in such a way that the strategies yielding the highest forgone payoffs continue to survive in the set of remembered strategies while other strategies are discarded. The foregone payoff is the payoff that a given strategy j, $j \in 1, ..., J$, would have received given what an opposing player actually played.

To apply IEL to the extended games studied in the experiments on correlated equilibria, the IEL players are provided with third party recommendations in the same manner as was done in the experiments with human subjects. Specifically, IEL agents are given a recommendation to play either action C or D. However, unlike the experimental subjects, IEL players are not given information on the probability distribution of recommendations nor are IEL players informed that other IEL players also receive recommendations, as IEL agents do not condition their behavior on this type of information. The first part of player i's strategy j, $s^i_{j,t}$ at t is $f^i_{j,t} \in [0, 1]$. This part represents the probability that player i follows the recommendation made by the third party in period t and is thus referred to as the *follow* part of the strategy. The second part of the strategy, $nf^i_{j,t}$ (for *not follow*), prescribes the action that player i takes whenever she does not follow the recommendation in period t. Each element of nf can take a value of 0 or 1 and will be referred to as *the action* part of the strategy. The remembered strategies of agent i thus take on the following form:

$$S^i_t = s^i_{1,t}, s^i_{2,t} \cdots s^i_{J,t}$$

with $s^i_{j,t} = (f^i_{j,t}, nf^i_{j,t}) \forall j = 1, \ldots, J$. The *follow* part of the strategy consists of a single element and the *not follow* part consists of two elements (thus, we use lengths $\ell(f) = 1$ and $\ell(nf) = 2$ in the simulations). The single element of f encodes the probability of following recommendations. The first element of nf determines the action the player takes in odd-numbered periods while the second element of nf determines the action the player takes in even-numbered periods.[2] The choice of two elements for nf is based on experiments showing that players using turn taking or alternating strategies, such as Duffy et al. (2016), Arifovic and Ledyard (2015), and Kuzmics et al. (2014), that would be impossible to achieve whenever $\ell(nf) = 1$.

[1] For example, see Arifovic and Ledyard (2007) for application to call markets; Arifovic and Ledyard (2010) for application in the Groves-Ledyard mechanisms for the allocation and financing of public goods; Arifovic and Ledyard (2012) for application to voluntary contribution games.

[2] This even-odd period structure is without loss of generality as the actions actually played in even or odd numbered periods are endogenously determined by the algorithm.

The initial set of remembered strategies, $S_1^i, \forall i$, are drawn randomly.[3] At $t = 1$, each strategy has an equal $(1/J)$ chance of being selected to be played. The set of strategies, S_t^i, co-evolves over time based on the recommendation received and the actions of other players. The updating of each player's strategy set takes place via *experimentation*, computation of *foregone payoffs*, *replication*, and *selection*.

The sequence of steps in our evolutionary algorithm is as follows:

Experimentation introduces new elements to the strategies in a player's strategy set. Let ρ denote the probability that any element of a strategy will undergo experimentation with the probability being independent across each element in a given strategy.

If the element of the *follow* part of the strategy (which consists of a single element) is selected, then a new element is drawn from a normal distribution with a mean equal to the value of the existing element, $f \sim N(f_{j,t}^i, \theta)$, $f \in [0, 1]$ otherwise f is redrawn.

If an element from the *nf* part of the strategy is selected for experimentation, the existing value may be flipped, i.e., $0 \to 1$ or $1 \to 0$. The probability of both experimentation and of being flipped for the *nf* part of the strategy is thus equal to $\rho/2$.

Computation of **Foregone Payoffs** determines the expected foregone payoff that the player could have achieved from a strategy $s_{j,t}^i \in S_t^i$ taking player i's history, h_t^i, as given. The history, h_t^i, consists of the observed actions of other players and the recommendations of play from the two most recently played periods. Allowing the IEL to account for correlated strategies (recommendations) requires us to take into account both the probability to follow or not follow the recommendation as well as the choice of pure actions in the computation of the foregone payoff. Thus, the expected foregone payoff is given by:

$$W(s_{j,t}^i|h_t^i) = f_j^i \sum_{k=0}^{1} u(r_{t-k}^i, \sigma_{t-k}^{-i}) + (1 - f_j^i) \sum_{k=0}^{1} u((nf)_{j,t-k}^i, \sigma_{t-k}^{-i}) \tag{3.1}$$

where $W(s_{j,t}^i|h_t^i)$ is the expected foregone payoff, r_{t-k}^i is the recommendation at $t - k$, and σ_{t-k}^{-i} is the action of player $-i$ at $t - k$.

Equation 3.1 consists of two parts: The first part is the payoff that strategy $s_{j,t}^i$ would have received from following the recommendation, weighted by the probability of following the recommendation, $f_{j,t}^i$, in periods t and $t - 1$. The second part is the payoff that the strategy $s_{j,t}^i$ would have received from the *nf* elements of the strategy in periods t and $t-1$, weighted by the probability of not following, $(1 - f_{j,t}^i)$.

Replication reinforces strategies that would have worked well in previous periods. The choice as to which strategies remain in the set of remembered strategies depends on the value of the expected forgone payoff. For a given set of values $W(s_{j,t}^i|h_t^i) = W_{j,t}^i$, two strategies $s_{l,t}^i$ and $s_{k,t}^i \in S_t^i$ are chosen randomly (with uniform probability and replacement) to compete

[3] Initialization: At the beginning of each run, all *follow* parts of each player's set of strategies are drawn from a uniform distribution, in the interval $[0, 1]$. All *not follow* parts of each player's set of strategies are assigned a value of 0 or 1 with equal probability.

for a chance to remain in the remembered set of strategies. The expected foregone payoffs of the two strategies are compared, and the one with the higher foregone payoff is selected.[4]

$$s^i_{j,t+1} = \left\{ \begin{array}{c} s^i_{l,t} \\ s^i_{k,t} \end{array} \right\} \text{if} \left\{ \begin{array}{c} W^i_{l.t} \geq W^i_{k,t} \\ W^i_{k,t} < W^i_{l,t} \end{array} \right\}.$$

The above steps are repeated J times. This completes the updating of the sets of strategies in period $t+1$.[5]

Selection probabilistically chooses a strategy that the player actually uses for the next two periods. A strategy $s^i_{j,t+1} \in S^i_{t+1}$ is chosen with probability $q^i_{j,t+1} = W^i_{j,t}/(\sum_{k=1}^{J} W^i_{k,t})$. [6]

In terms of parameters, we have already discussed how strategies are represented as two component structures with $\ell(f) = 1$ and $\ell(nf) = 2$. Given the strategy representation, the IEL model has three free parameters (J, ρ, θ). We chose $(J, \rho, \theta) = (180, 0.033, 0.10)$ as these values have proven to work well in capturing behavior in a number of different experiments with human subjects (for example, Arifovic and Ledyard (2007, 2010, 2012) and Boitnott (2015)). Thus, the strategy set of player i, at period t, is given by:

$$S^i_t = \begin{bmatrix} s_{1,t} \\ s_{2,t} \\ \vdots \\ s_{180,t} \end{bmatrix} = \begin{bmatrix} f_1 & nf_{1,1} & nf_{1,2} \\ f_2 & nf_{2,1} & nf_{2,2} \\ \vdots & \vdots & \vdots \\ f_{180} & nf_{180,1} & nf_{180,2} \end{bmatrix}$$

Later, in section 3.5, we will consider a grid search to determine whether the parameterization of the IEL model could be changed to better fit the data.

The design of our IEL environment and simulations closely follow the experimental design and treatments of Duffy et al. (2016) for the Battle of the Sexes game and Duffy and Feltovich (2010) for the game of Chicken. We conducted 500 simulation runs of IEL for each *treatment* of these two experimental studies. We next describe these treatments.

3.4.1 Battle of the Sexes

Duffy et al. (2016) (henceforth DLL) study the stage game (Table 3.1) played repeatedly for 60 periods. Their experiment utilizes two matching protocols and three treatments regarding recommendations. The matching protocols have players paired using either a Partners (fixed matching) or a Strangers (random matching) design. For the Partners matching design, two subjects are matched at the beginning of a game and remain in the

[4]In case the two payoffs are equal, a coin flip is used to select a strategy.

[5]While it is customary in the literature on evolutionary learning to implement replication before experimentation, we use the reverse sequence in order to increase the speed at which the learning takes place. However, the qualitative features of the IEL's behavior do not depend on a sequence in which replication and experimentation take place.

[6]If some values of $W^i < 0$, then the probability is computed as $q^i_{j,t+1} = (W^i_{j,t} - \epsilon^i_t)/\sum_{k=1}^{J}(W^i_{k,t} - \epsilon^i_t)$ where $\epsilon^i_t = min\{0, W^i_{1,t}, \ldots, W^i_{J,t}\}$ for use in period $t+1$.

same fixed pairing for all 60 periods of the game. Under the Strangers matching protocol, six subjects are divided up equally between row and column players and in each period, three pairs are formed by randomly matching together one row and one column player.

Both matching protocols are implemented in their three main treatments. In their baseline, "direct" recommendations treatment, prior to choosing an action in each period, subjects are given a recommendation to play either action C or D with the knowledge that both players in a pair would receive the *same* recommendation, r, and that $\Pr(r = C) = \Pr(r = D) = .5$. In their second, "None" treatment, subjects are not given any recommendation as to how to play the game. In their third, "indirect" recommendations treatment, subjects are given recommendations that do not directly map into the action space; in that treatment, the recommendations are either @ or # to both players with $\Pr(r = @) = \Pr(r = \#) = .5$. The latter treatment allows for the development of an endogenous mapping between recommendations and the action space of the stage game. In order to examine how well the behavior of IEL matches the behavior observed in the experiment with human subjects, we use the following protocol in our IEL simulations.

We match our simulation design to the design of DLL's experiment with six different treatments: {direct recommendations, no recommendations (none) and indirect recommendations} $\times$ {Partners, Strangers}. Thus, the IEL players are given either a direct recommendation to play either C or D, no recommendation or an indirect recommendation either @ or #. We conducted 500 simulation runs of the IEL algorithm for each of the six treatments with a length of each run equal to 60 periods, the same number of periods in DLL's experiment.

Following DLL, we split the IEL players into two groups: *row* players and *column* players with a row player being matched with a column player using one of the two matching protocols. Under Partners (Fixed) matching, we have two IEL players interact with one another for 60 periods. In the Strangers (random) matching, we have six IEL players interact together for the 60 periods with 3 pairs being randomly formed in each period.

For each pair of players, a period proceeds as follows. First, in treatments with recommendation, a recommendation for each pair of players is determined by drawing a random number from a uniform distribution over $[0, 1]$. If the number is less or equal to 0.5, the recommendation given to the pair is to play $C(@)$ in that period. Otherwise, the recommendation given to the pair is to play $D(\#)$ in that period. We next determine whether players follow this recommendation. For each player, a random number, ϕ, is drawn from a uniform distribution over $[0, 1]$. If ϕ is less than or equal to the probability that the recommendation is followed as given in $f^i_{j,t}$, then the player follows that recommendation.[7] Otherwise, if recommendations are not followed, or if no recommendations are given (as in the 'None' treatment), the player uses the two-element $nf^i_{j,t}$ part of her strategy to determine what

[7]In the indirect recommendation treatment, each player must also assign a meaning to the signal, as discussed in further detail below.

strategy to play depending on whether it is an odd- or even- numbered period. If the value of the nf element corresponding to the period (odd or even) is 0, then the player plays C, while if the value is 1 she plays D.

As mentioned earlier, a player uses a selected strategy for two consecutive periods. After every two periods, expected forgone payoffs are computed for all the strategies in each player's remembered strategy set, S_t^i, taking the observed history as given (i.e., the actions of the other player and the recommendations over the previous two periods). The forgone payoffs are then used to create a new set of remembered strategies, S_{t+1}^i, and to determine via selection the strategy to be played during the next two periods.[8]

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.911 (0.092) [0.841,0.982]	5.47 (0.548) [5.04,5.89]	0.993 (0.015) [0.982,1.00]	5.96 (0.088) [5.89,6.02]
Sim. Strangers	0.872 (0.049) [0.868,0.876]	5.23 (0.295) [5.20,5.26]	0.976 (0.049) [0.971,0.980]	5.85 (0.293) [5.83,5.88]
DLL Partners	0.857 (0.169) [0.727,0.988]	5.14 (1.016) [4.36,5.93]	0.878 (0.331) [0.624,1.132]	5.27 (1.886) [3.74,6.79]
Sim. Partners	0.863 (0.083) [0.856,0.870]	5.18 (0.496) [5.13,5.22]	0.963 (0.088) [0.956,0.971]	5.78 (0.528) [5.73,5.83]

Table 3.3: BoS Direct Recommendations Comparison

Table 3.3 shows simulation results for the Battle of the Sexes (BoS) Direct Recommendations (DR) treatment under both the Strangers and Partners matching protocols in comparison with the human subject outcomes reported in DLL. Comparisons are made with respect to the mean frequencies of coordination on pure strategy equilibria and standard deviations (in parentheses) over two time intervals and with the mean and standard deviation of payoffs actually earned over the same two time intervals.[9] In addition, we provide 95 percent confidence intervals [in square brackets] for coordination rates and payoffs for both the human subject and simulated data based on the standard deviations from those data. Recall that the IEL simulation means and standard deviations are calculated from 500 runs.

The results in Table 3.3 suggest that there is a reasonably good fit between the means of the simulation runs and the human subject coordination and payoff outcomes; overall 60-period simulation means lie within the 95 percent confidence bounds of the data, though the

[8]Later, in section 3.5, we explore the sensitivity of our findings to expanding the number of periods used to evaluate foregone payoffs beyond just 2.

[9] The payoff means and standard deviations from the 500 simulation runs (Sim.) reported on in Table 3.3 (and later for Tables 3.5 and 3.7) are *actual* payoffs earned by the simulated strategies and not foregone payoffs. While the IEL uses foregone payoffs to evaluate the fitness of strategies according to equation (1), these foregone payoffs are not the same as the payoffs actually earned from play of the chosen strategies and it is the latter that we are reporting on here for comparison with the payoffs actually earned by DLL's human subject participants.

standard deviations in the humans subject data are generally greater than in the simulated data. Perhaps the most important finding from Table 3.3 is that in both the DLL experimental data and the IEL simulations, coordination frequencies and payoffs are higher under the Strangers matching protocol as compared with the Partners matching protocol. This finding is notable because the IEL algorithm does not explicitly condition on the matching protocol while human subjects were aware of whether they were operating under a Partners or Strangers matching scheme. The reason for this difference, as explained in DLL, is that under the partners matching protocol, subjects have an alternative means of coordination, namely using their history of fixed interactions, and this reduces the importance of the direct recommendations. However, in the Strangers matching protocol, the recommendations are more useful for solving the coordination problems, and generate higher payoffs as well.

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
DR	T = 1-60	0.0%	0.0%	0.0%	0.0%	0.85
Strangers	31-60	0.5%	0.3%	0.4%	0.3%	0.96
	46-60	1.2%	0.8%	0.8%	0.8%	0.97
DR	T = 1-60	1.3%	1.0%	0.7%	0.3%	0.76
Partners	31-60	4.1%	3.7%	4.0%	3.6%	0.83
	46-60	4.3%	4.2%	4.3%	4.3%	0.84

Table 3.4: IEL nf and f Strategies in BoS-DR

Table 3.4, reports on the strategies being played by the IEL simulated agents in the BoS-DR simulations under the Strangers or Partners matching protocols and over three sample periods, T=1-60, 31-60 or 46-60. Recall that the nf part of the strategy has two components, $nf^i_{j,1}$, $nf^i_{j,2}$, each of which can either be C (0) or D (1), so there are just four possible outcomes for the nf part of a strategy. We classify a simulated player as having the not follow strategy of $[nf^i_{j,1}, nf^i_{j,2}]$, if at least 90% of his strategy set of $J = 180$ strategies (at least 162 strategies) have the *same* bit encodings for $[nf^i_{j,1}, nf^i_{j,2}]$, over the sample period, e.g., a nf strategy of [C D] means that at least 90 percent of the agent's strategies have a nf part that specifies the play of C in odd periods and D in even periods over the sample period. The threshold of 90 percent was chosen to reduce noise in the determination of the nf part of agents' strategies. We also report on the average value of the follow part, f, of agents' strategies over the same three sample periods. Recall that f is simply a real number in $[0, 1]$, representing the probability that the simulated agent chooses to follow recommendations; here we simply report the average value of f over the given sample period. As Table 3.4 reveals, by the end of both the Strangers and Partners BoS-DR treatments, the average value for f is .97 and .84, respectively, indicating a high likelihood of recommendation following in both treatments. Consequently, there is not much consistency in the not follow (nf) parts of simulated agents' strategies, as these strategies are not being played very much, and this observation accounts for the low percentage of nf strategies meeting our 90 percent threshold for classification.

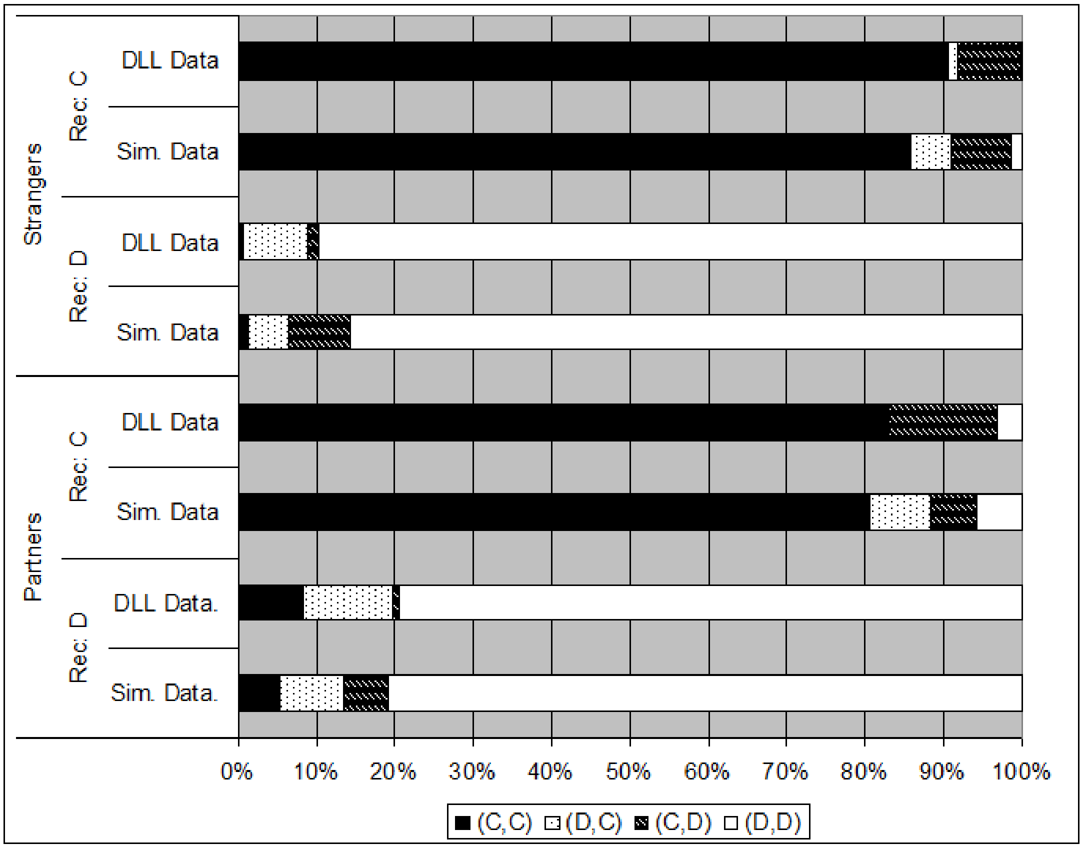
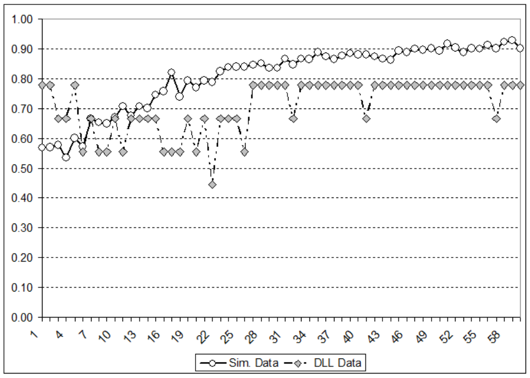


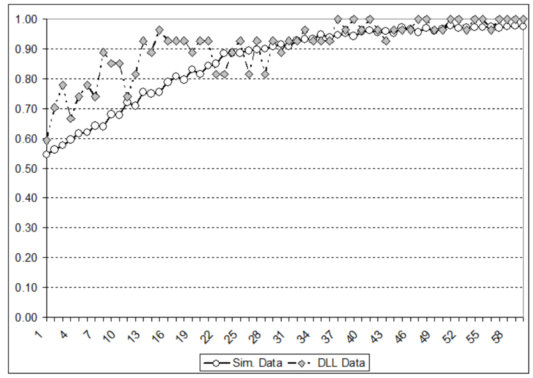
Figure 3.1: Outcome Frequencies Based on Recommendations in BoS-DR

Figure 3.1 provides a more disaggregated comparison of the mean frequencies of actual outcomes observed in DLL's DR treatment and in the IEL simulations based on the recommendations that were given, with 'Rec: C' and 'Rec: D' indicating which action was recommended under the Strangers or Partners matching protocols. Thus for instance, we observe that when both players were recommended to play C (D) they did so with a high frequency in the Partners, DR treatment and with an even higher frequency in the Strangers, DR treatment, consistent with the DLL experimental data.

Figure 3.2 shows the frequency with which *pairs* of players followed recommendations over time in both the IEL simulations and DLL data. The time series shown are the frequencies with which both members of each pair of players followed the recommendations given to them and thus this frequency differs from the average value of the follow part of all players' strategies, as reported on in Table 3.4. We compare our simulated data with the experimental data using this pairwise recommendation following frequency as we have such data for both of the experimental studies (BoS and Chicken) with which we compare our simulation results. Notice that, consistent with the experimental data, recommendation



(a) BoS-DR Partners



(b) BoS-DR Strangers

Figure 3.2: Pairwise Recommendation Following over Time in the BoS-DR treatments: Simulations versus DLL Data

following increases over time in both the Partners and Strangers versions of the BoS-DR treatment, though the IEL simulations provide a better fit to the DLL-Strangers data than to the DLL-Partners data.

Simulation results for the treatment where no recommendations are given – the "None" treatment of DLL – as compared with the corresponding experimental data of DLL are reported in Table 3.5. The simulation results for this case are obtained by setting $f_j^i = 0\,\forall j, i$ and also setting the probability of experimentation with this part of each strategy equal to zero. These changes ensure that all players are classified as not following recommendations as reflected in Table 3.6. Reported results are again from 500 simulations, and, as in Table 3.3, Table 3.5 reports means, (standard deviations) and [95 percent confidence intervals] for coordination rates and payoffs over two sample periods. In addition, for the two versions of None, the aggregate outcome frequencies are reported in Figure 3.3 along with the corresponding experimental outcome frequencies from DLL. The simulated and human subject data are reasonably consistent with the notable exception of the frequencies of (C,D) and

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.548 (0.141) [0.439,0.656]	3.29 (0.843) [2.64,3.93]	0.596 (0.203) [0.440,0.753]	3.58 (1.218) [2.64,4.51]
Sim. Strangers	0.836 (0.075) [0.829,0.842]	5.01 (0.449) [4.97,5.05]	0.980 (0.064) [0.974,0.985]	5.88 (0.387) [5.84,5.91]
DLL Partners	0.869 (0.185) [0.726,1.011]	5.21 (1.111) [4.36,6.06]	0.800 (0.316) [0.557,1.043]	4.80 (1.897) [3.34,6.26]
Sim. Partners	0.899 (0.071) [0.893,0.906]	5.40 (0.425) [5.36,5.43]	0.987 (0.037) [0.984,0.990]	5.92 (0.219) [5.90,5.94]

Table 3.5: BoS No Recommendations (None) Comparison

(D,D) which are too low and too high (respectively) in the simulated Strangers treatment relative to comparable DLL Strangers data.[10]

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
None	T = 1-60	2.5%	2.1%	2.0%	2.5%	0.00
Strangers	31-60	19.9%	21.5%	21.8%	23.8%	0.00
	46-60	21.8%	23.0%	23.8%	26.1%	0.00
None	T = 1-60	6.6%	8.7%	7.1%	5.9%	0.00
Partners	31-60	25.1%	23.0%	25.2%	26.0%	0.00
	46-60	25.2%	23.4%	25.2%	26.2%	0.00

Table 3.6: IEL nf and f Strategies in BoS-None

Table 3.6 again reports on the percentage of the nf parts of the simulated strategies that meet the 90 percent threshold for strategy classification as discussed earlier in connection with Table 3.4. Here we see that, since the f part is constrained to be zero, the nf parts of the simulated agent strategies settle on play of the two pure strategies, [D D] or [C C], or on alternation strategies, [D C] or [C D], and by the last 15 periods, 94.7 or 100 percent of simulated agents (Partners, Strangers, respectively) meet our nf strategy clarification criterion.

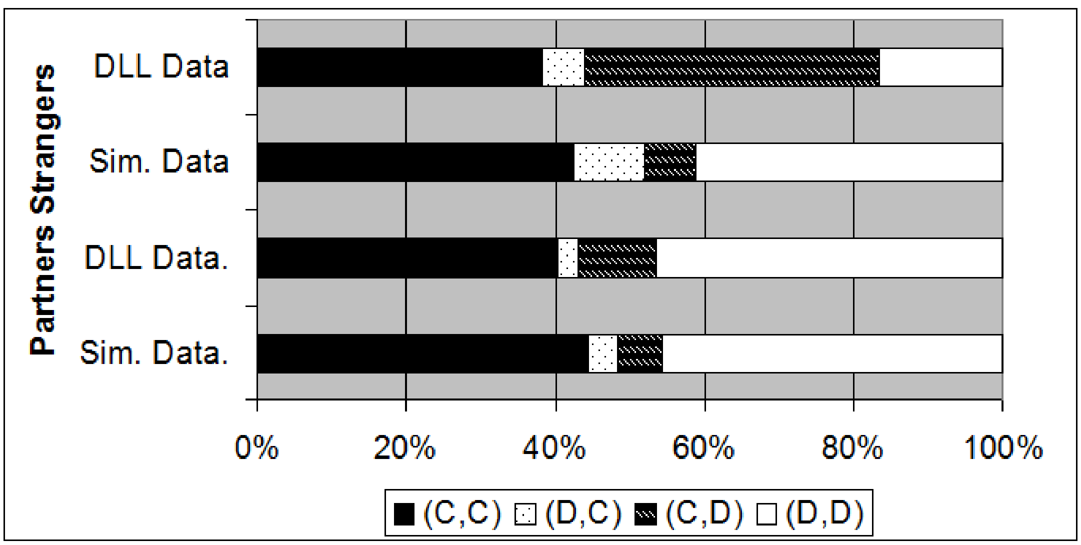


Figure 3.3: Observed Outcome Frequencies in BoS-None

Finally, we report results for the BoS Indirect Recommendations or IR treatment. This environment is more complicated to implement algorithmically as the mapping from the randomization device to the action space is no longer clearly specified since the signal, or

[10]As shown in Table 6 of DLL, in the None-Strangers treatment, several of the human subject groups coordinated on the unique mixed strategy Nash equilibrium (MSNE) of the game which would result in the (C,D) outcome being played 56.25% of the time. The IEL algorithm, with its 2 period history, remains more likely to learn a pure strategy or alternation strategy equilibrium than play of the MSNE.

#, does not directly map into the two actions, C or D. Thus, additional assumptions are needed. Our approach is to assume that there are 4 possible interpretations of the signal:

1. @ signals the player should play C and # signals the player should play D
2. # signals the player should play C and @ signals the player should play D
3. @ and # signal the player should play C
4. @ and # signal the player should play D

For simplicity, we ignore outcomes 3 and 4 from the list above and focus on strategies that interpret different signals as having a unique meaning since the strategy representation is already capable of ignoring recommendation/signals and playing a pure strategy. Thus for the IR environment, each player has an extra element added to their strategy, λ, which is initially a uniform random draw over $[0,1]$. This new part, λ, (like the follow part, f) represents the probability that @ signals play of action C while # signals play of action D and $1-\lambda$ represents the opposite mapping. The strategies including λ are evaluated using the same method previously discussed with the exception that the player may not know the precise information given by the signal when choosing whether to follow the signal or not. This means the player knows what strategy they played (i.e., C or D), the payoff they received, and the strategy played by the other player, but not whether $\# = C$ or $@ = C$. The expected forgone payoff then becomes:

$$\begin{aligned} W(s^i_{j,t};h^i_t) =& f^i_j(\lambda^i_j((\sum_{k=1}^{2} u(\kappa^i_{t-3+k},\sigma^{-i}_{t-3+k}))/2) + (1-\lambda^i_j)(\sum_{k=1}^{2} u(\varrho^i_{t-3+k},\sigma^{-i}_{t-3+k}))/2) \\ &+ (1-f^i_j)(\sum_{k=1}^{2} u((nf)^i_{j,k},\sigma^{-i}_{t-3+k}))/2 \end{aligned} \tag{3.2}$$

where $\kappa^i_t = (C|@, D|\#)$ and $\varrho^i_t = (D|@, C|\#)$.

Table 3.7 reports means, (standard deviations) and [95 percent confidence intervals] for coordination rates and payoffs over two subsamples of the BoS-IR treatment for both the human subject DLL data and the IEL simulations. Consistent with DLL's experimental data, we observe that the IEL algorithm learns to coordinate more effectively and with higher payoffs under the fixed Partners matching protocol as compared with the random Strangers matching protocol in the IR treatment. The IEL's strategies for the IR treatment are reported on in Table 3.8 which uses the strategy classification discussed earlier in connection with Table 3.4. This table reveals that the average probability of recommendation following, f, is .30 for the last quarter of the Strangers treatment but is decreasing more rapidly toward zero in the Partners treatment. Indeed, the simulated agents in this IR treatment are learning to ignore recommendations, and more rapidly so under the partners matching protocol, as they instead choose to play pure [C C],[D D] or alternation [C,D] [D,C] strategies that do not follow recommendations. Additionally, while many of the agents are shown to eventually coordinate on a single strategy, the agents perform worse than the case where no recommendations are given. Figure 3.4 provides a break down of the frequency of outcomes given the signal. The simulations reveal a fairly consistent frequency of coordination, regardless of the recommendations given, whereas the experimental subjects appear to have been influenced by the recommendations to coordinate on a particular outcome especially under the Partners matching protocol.

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.580 (0.119) [0.488,0.671]	3.48 (0.715) [2.93,4.03]	0.674 (0.174) [0.532,0.816]	4.04 (1.108) [3.19,4.90]
Sim. Strangers	0.608 (0.086) [0.600,0.615]	3.65 (0.516) [3.60,3.69]	0.735 (0.162) [0.721,0.749]	4.41 (0.972) [4.32,4.49]
DLL Partners	0.915 (0.128) [0.808,1.022]	5.49 (0.768) [4.84,6.13]	1.000 (0.000) [1.000,1.000]	6.00 (0.000) [6.00,6.00]
Sim. Partners	0.761 (0.133) [0.749,0.773]	4.57 (0.800) [4.50,4.64]	0.934 (0.140) [0.922,0.946]	5.60 (0.842) [5.53,5.68]

Table 3.7: BoS Indirect Recommendations Comparison

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
IR.	T = 1-60	0.0%	0.1%	0.0%	0.0%	0.39
Strangers	31-60	9.9%	8.5%	8.9%	7.1%	0.33
	46-60	14.0%	11.9%	13.5%	11.4%	0.30
IR.	T = 1-60	1.1%	2.4%	0.7%	0.4%	0.26
Partners	31-60	22.5%	26.5%	19.9%	15.6%	0.13
	46-60	25.1%	28.5%	20.9%	17.0%	0.09

Table 3.8: IEL nf and f Strategies in BoS-IR

Finally, Figure 3.5 shows the frequency of pairwise recommendation following over time in the BoS-IR treatments. The IEL simulations clearly underpredict the extent of recommendation following in the Partners version of the BoS-IR treatment, most likely because the algorithm does not exploit the fact that players are repeatedly rematched with the same other player. However, for the strangers BoS-IR treatment, the fit to the frequency of recommendation following over time is very good, with the simulated data tracking the DLL data very closely albeit with less noise. Note further that the IEL also gets right the initial frequency of recommendation following in both BoS-IR treatments, which is only around 25 percent and not the 50 percent level observed in BoS-DR treatment. The reason for this difference is that in the IR treatment, the mapping from the recommendations to the action space, (λ) also has to be learned, and so initially, only about 50 percent of the strategies have the right mapping. Combining this fact with the 50 percent initial chance of following

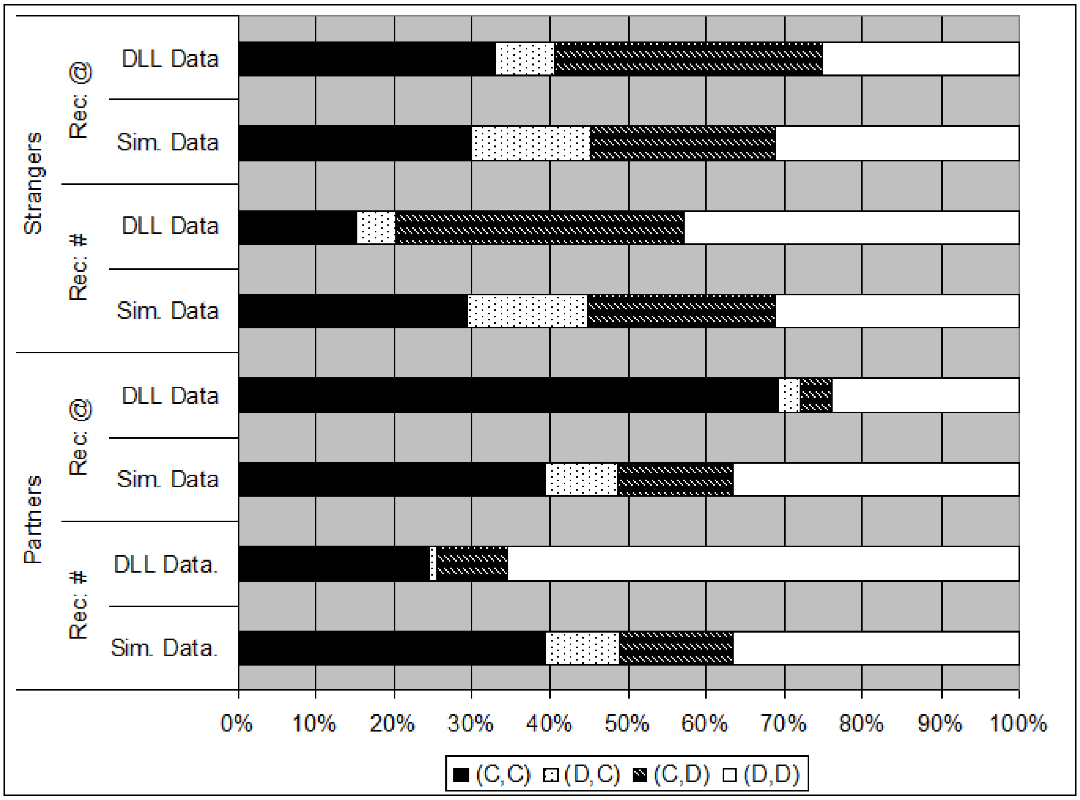


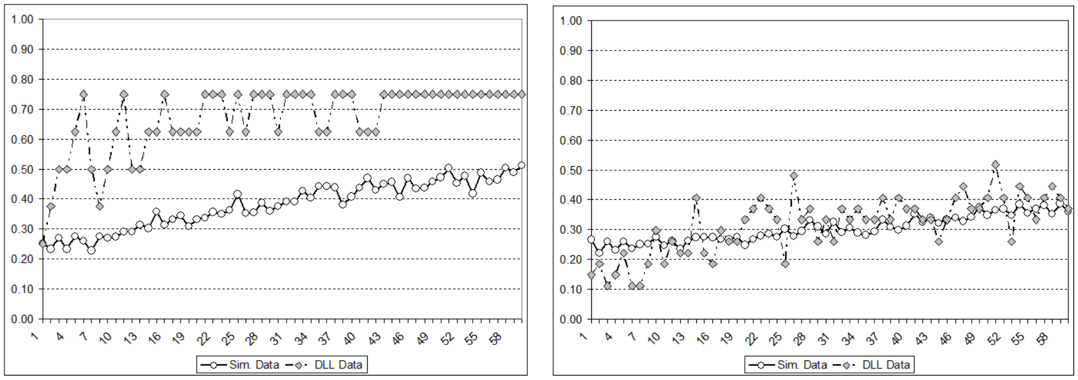
Figure 3.4: Outcome Frequencies based on Recommendations in BoS-IR

recommendations yields an approximately 25 percent chance of initial recommendation following, which also shows up in DLL's experimental data. Overall, we conclude that there is a good qualitative fit between the IEL simulated data and DLL's experimental data for the BoS game.

3.4.2 Chicken

We next address the learning of correlated equilibria in the Chicken game, which is a more complicated coordination game than BoS, and which allows for some more interesting correlated equilibria that lie outside the convex hull of Nash equilibrium payoffs. We again implement our IEL simulations for the Chicken game in a manner that closely emulates the design of the experiments with human subjects as reported in Duffy and Feltovich, 2010, (DF hereafter), so as to facilitate comparisons. The stage game is shown in Table 3.2.

The DF experiments involved five treatments using four different sets of recommendations, r. The different recommendations consist of three correlated equilibria called the *Nash*, the *Good*, and *Bad* recommendations equilibria. The *Nash* recommendations treatment has $\Pr(r^i = D | r^{-i} = C) = \Pr(r^i = C | r^{-i} = D) = .5$, where r^{-i} is the recommendation given to the other player. This treatment most closely aligns with the recommendation used in the Battle of the Sexes game, but here, if the row player is recommended to play C, the



(a) BoS-IR Partners

(b) BoS-IR Strangers

Figure 3.5: Pairwise Recommendation Following over Time in the BoS-IR treatments: Simulations versus DLL Data

column player is recommended to play D and vice versa. As in BoS, this correlated equilibrium is an equal mixture between the two pure strategy Nash equilibria, hence the name 'Nash' correlated equilibrium. The next correlated equilibrium is induced using probabilities $\Pr(r^i = D|r^{-i} = C) = \Pr(r^i = C|r^{-i} = C) = \Pr(r^i = C|r^{-i} = D) = 1/3$ and is referred to as the *Good* recommendations equilibrium; this corresponds to the correlated equilibrium discussed earlier where payoffs lie outside the convex hull of payoffs that are possible from randomization over the set of pure strategy Nash equilibria and are greater than payoffs earned in the unique mixed strategy Nash equilibrium. The final correlated equilibrium is referred to as the *Bad* recommendations equilibrium with each player receiving a recommendation r with $\Pr(r^i = C|r^{-i} = D) = \Pr(r^i = D|r^{-i} = C) = 0.4$ and $\Pr(r^i = D|r^{-i} = D) = 0.2$; this equilibrium yields a payoff that is worse than the mixed strategy Nash equilibrium, hence the name 'Bad'. The last recommendation treatment is called the *Very Good* recommendations treatment since following the recommendations in that treatment yields a payoff that is better than any payoff from the set of possible correlated equilibria; however, since following the Very Good recommendations is *not* part of a correlated equilibrium – there are profitable deviations from following recommended play – the recommendations in the Very Good treatment should not be followed. Hence, this treatment serves as a check on whether players are blindly following recommendations or not. The Very Good treatment has recommendations of $\Pr(r^i = D|r^{-i} = C) = \Pr(r^i = C|r^{-i} = D) = 0.1$, and $\Pr(r^i = C|r^{-i} = C) = 0.8$. The experiment also includes results for the case where no recommendations are given referred to as the *None* treatment. DF only use the Strangers or random matching protocol in all of their treatments; unlike in the BoS game, there are no treatments using a Partners matching protocol in Chicken.

Thus, there are a total of 5 experimental treatments in the Chicken game to study using IEL. Our simulation uses the same Strangers matching protocol as in DF but with the number of computational subjects in a matching group increased from 6 to 12 to match the human subject experimental design. The simulations have the IEL agents play the game for 20 periods which is the same number of periods for which the experimental subjects in DF received recommendations. For each of the five treatments, we again report on 500 runs of the IEL simulation. The parameterization of the IEL model, $(J, \rho, \theta) = (180, 0.033, 0.10)$ remains the same in these Chicken game simulations as was used in the BoS simulations.[11]

There is an important difference between what is done in the experimental setting and the simulations that requires some further discussion. The experiment consists of two separate parts. Each part consists of the human subjects playing for 20 periods. No recommendations are provided during the first (or second) part of the experiment and one of the four recommendations treatment conditions is in place during the second (or first) part. Because we are interested in determining if and how the IEL agents are able to learn correlated equilibrium, our simulations are run without this feature. Although including both parts of the experiment may have improved how closely the algorithm matches the experimental data, it would have muddled our understanding of the outcomes generated by the algorithm.

Table 3.9 and Figure 3.6 provide a comparison between the experimental data in the various DF Chicken game treatments and the IEL simulations. We note that each recommendation structure provides different expected payoffs. In order to keep a consistent comparison of the payoffs, we choose to make comparisons against the expected payoff from using the mixed strategy Nash equilibrium, which is 5.4. In other words, our "efficiency measure" as reported on in Table 3.9, is $100 * \sum_{i=1}^{I} \sum_{t=1}^{T} \pi_t^i / (I * T * 5.4)$ or the average payoff over $T = 20$ periods and across I agents divided by the mixed strategy Nash Equilibrium payoff. As Table 3.9 reveals, consistent with the experimental data, payoffs earned (efficiency scores) by the simulated agents are higher in the Good and Nash treatments as compared with the Bad or None treatments. However, for the Very Good treatment, the IEL follows recommendations too closely, which leads to higher payoffs, but the human subjects learn more quickly to ignore recommendations in this treatment so their payoffs (efficiency scores) are lower. Nevertheless the outcome frequencies in the simulated data are not far off from the experimental outcome frequencies as Figure 3.6 reveals.

Source:	Good	Nash	Bad	Very Good	None
DF Data	100.81	104.59	95.87	99.70	99.76
Sim. Data	101.76	101.40	91.77	104.53	95.82

Table 3.9: Comparison of Efficiency Measures in all Chicken Treatments

[11] Later in section 5, we report on a grid search exercise to find the parameterization of IEL that is a best fit to the experimental data which we conduct separately for Chicken and BoS.

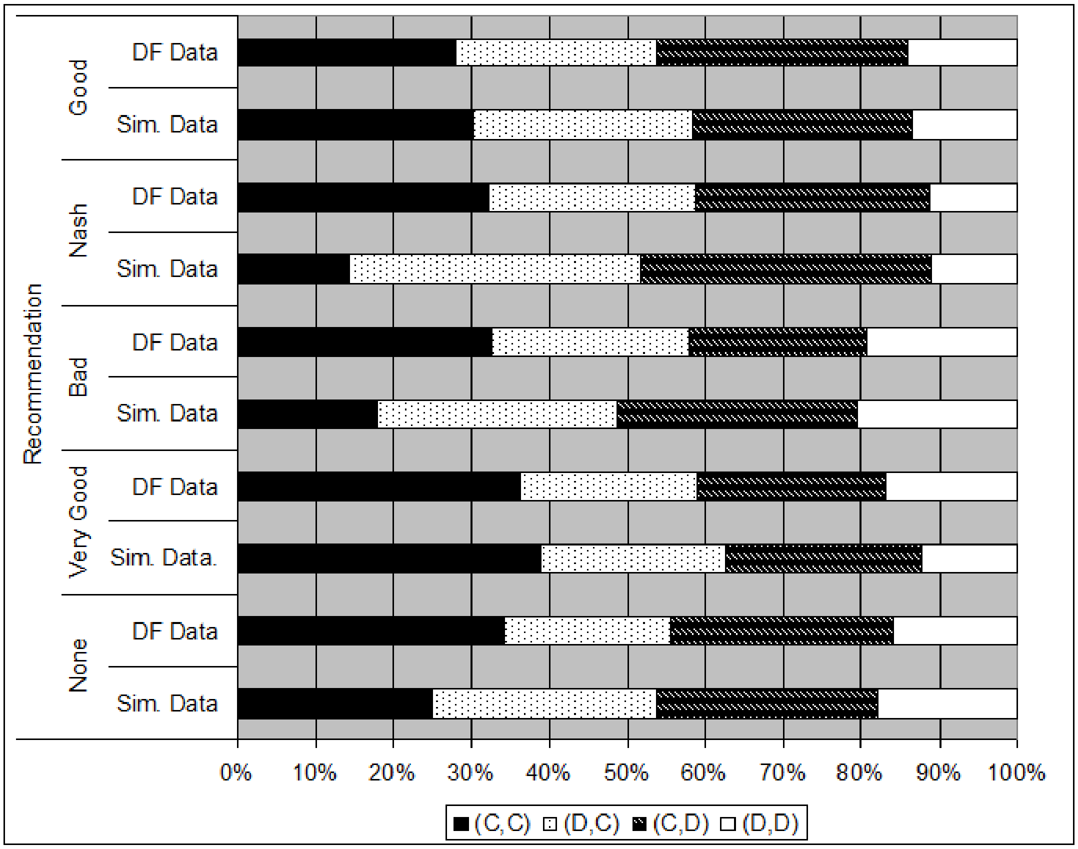


Figure 3.6: Observed Outcome Frequencies in all Chicken Treatments

Table 3.10 reports on various frequencies of recommendation following in the DF data and IEL simulations. The individual frequency of recommendation following regardless of the recommendation received appears as "Total" but this frequency is further broken down according to whether the recommendation was C or D. The frequency of *pairwise* recommendation following (as studied in the BoS game) is labeled "Pairwise". We report on all of these different measures of recommendation following as we have data on these measures from DF's data set. In Table 3.10 the mean squared error (MSE) is based on the difference between the frequency of recommendation following in the DF and IEL data. We observe that, while these MSEs are generally small, they do not change much over time; there is often a persistent level difference in that the frequency of recommendation following is higher in the simulated data than in the corresponding experimental data. However, consistent with the experimental data, pairwise recommendation following in the simulated data increases over time in the Good, Nash and Bad treatments, while it decreases over time (as it should) in the Very Good recommendations treatment. Notice further that in three of the four recommendation treatments, the human subjects appear to be conditioning their recommendation following on the recommendation received, e.g., following a C recommendation much more frequently than a D recommendation. By contrast, the IEL algorithm

% Follow	Good Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	73.5%	79.6%	0.011	73.0%	78.4%	0.012	75.0%	83.3%	0.009
Rec = C	73.2%	77.5%	0.011	71.9%	77.0%	0.012	77.0%	79.2%	0.007
Total	73.3%	78.2%	0.009	72.3%	77.4%	0.010	76.2%	80.6%	0.004
Pairwise	53.1%	61.3%	0.021	51.4%	60.0%	0.025	58.3%	65.0%	0.009
% Follow	Nash Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	56.7%	82.7%	0.076	55.3%	80.0%	0.070	60.8%	90.7%	0.092
Rec = C	77.7%	85.9%	0.016	77.2%	83.7%	0.012	79.2%	92.9%	0.027
Total	67.2%	84.3%	0.033	66.3%	81.9%	0.027	70.0%	91.8%	0.049
Pairwise	45.4%	71.6%	0.074	43.3%	67.3%	0.062	51.7%	84.4%	0.108
% Follow	Bad Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	47.7%	74.8%	0.081	46.0%	74.1%	0.087	52.9%	77.0%	0.062
Rec = C	63.1%	84.0%	0.070	66.5%	82.6%	0.043	53.0%	88.0%	0.150
Total	54.1%	78.5%	0.064	54.5%	77.5%	0.056	52.9%	81.4%	0.084
Pairwise	26.9%	61.7%	0.127	25.9%	60.1%	0.124	30.0%	66.4%	0.136
% Follow	Very Good Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	51.1%	77.3%	0.134	60.6%	76.9%	0.059	22.7%	78.6%	0.342
Rec = C	60.8%	67.8%	0.009	63.2%	68.8%	0.008	53.7%	64.8%	0.014
Total	59.9%	68.8%	0.014	62.9%	69.6%	0.009	50.8%	66.2%	0.027
Pairwise	36.5%	46.8%	0.019	40.0%	48.3%	0.015	25.8%	42.2%	0.032

Table 3.10: Observed Frequencies of Recommendation Following in all Chicken Treatments

does not condition on recommendations, nor is the IEL algorithm even aware of the distribution of recommendations which *was* known to the human subjects. The IEL algorithm simply decides first whether to follow recommendations or not, and if so, it plays the recommended action. These different informational/strategic assumptions likely account for some of the observed differences between the human subject and simulated data. Still, it seems remarkable (at least to us) that the simple IEL algorithm is able to capture important qualitative features of the experimental data from DF's Chicken game.

Table 3.11 reports on the types of strategies the IEL algorithm coordinated on over time, using the strategy classification methodology discussed earlier in connection with Table 3.4. We see again that the average frequency of recommendation following (the average value of f) is increasing over time in all of the treatments where recommendation following comprises a correlated equilibrium (Good, Nash and Bad) while it is decreasing over time in the Very Good recommendation treatment, where following recommendations is not part of

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.58
Good	11-20	1.9%	0.3%	1.0%	1.2%	0.61
	16-20	3.9%	0.9%	2.0%	2.2%	0.63
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.69
Nash	11-20	0.5%	0.8%	0.5%	0.6%	0.79
	16-20	1.1%	1.9%	1.1%	1.3%	0.83
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.59
Bad	11-20	0.4%	1.8%	0.8%	0.8%	0.64
	16-20	0.8%	4.2%	1.4%	1.5%	0.66
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.45
Very Good	11-20	3.9%	0.5%	1.6%	1.6%	0.43
	16-20	7.1%	1.1%	3.4%	3.3%	0.42
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.00
None	11-20	3.8%	3.4%	2.9%	3.3%	0.00
	16-20	7.1%	6.5%	5.6%	6.5%	0.00

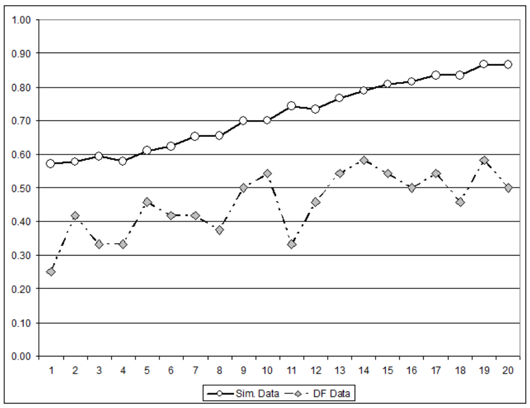
Table 3.11: IEL nf and f Strategies in all Chicken Treatments

a correlated equilibrium. This is reassuring, as it indicates that the algorithm is not blindly learning to follow recommendations, regardless of whether it is a best response to do so. When our simulations are allowed to run for a longer period of time than in the experiment– 100 periods– this pattern becomes even more pronounced with recommendation following increasing to very high levels in the Good, Nash or Bad recommendation treatments and falling further in the Very Good recommendations treatment.

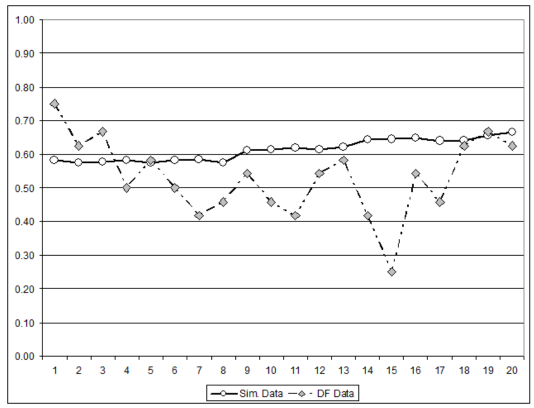
Finally, Figure 3.7 shows time series on the frequency of pairwise recommendation following over all 20 periods in the four Chicken treatments with recommendations: Nash, Good, Bad, and Very Good. As noted in the discussion of Tables 3.10-3.11, while there are important level differences between the frequency of pairwise recommendation following in the IEL simulations and the DF data, especially in the Chicken-Nash and Chicken-Bad treatments, the IEL algorithm generally gets the trend in pairwise recommendation following correct.

3.5 Sensitivity Analysis

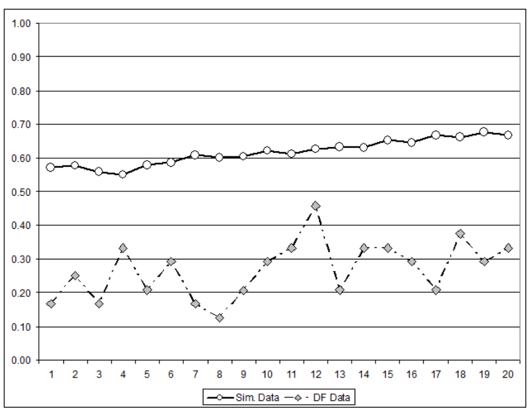
The IEL simulation results reported in the previous section provide a good, though not perfect characterization of the behavior of the experimental subjects in the Battle of the Sexes and Chicken games. In this section we report on a sensitivity analysis we conducted for our application of the IEL model to the learning of correlated equilibria in these two games. One aim was to check the robustness of our simulation results, but another goal was to see if we could improve the fit of our model to the experimental data. To address the robustness question, we consider the impact of modifications to both the model and its



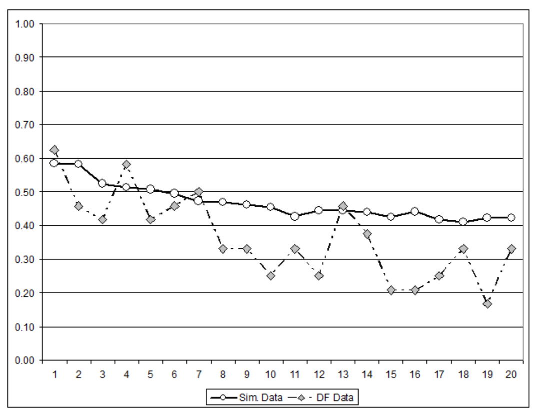
(a) Chicken-Nash



(b) Chicken-Good



(c) Chicken-Bad



(d) Chicken-Very Good

Figure 3.7: Pairwise Recommendation Following over Time in the Chicken Treatments: Simulations versus DF Data

initialization. We further conduct a grid search to evaluate whether changes in the three free parameters of the IEL model would improve the fit of the model to the experimental data, and we examine how sensitive our results are to small changes in a single IEL parameter from the "baseline" values chosen for the simulations reported in the previous section. Regarding fit, we mainly rely on the mean squared error (MSE) between the period-by-period mean payoffs earned by the experimental subjects and by the simulated players.[12]

3.5.1 Foregone Payoff Evaluation

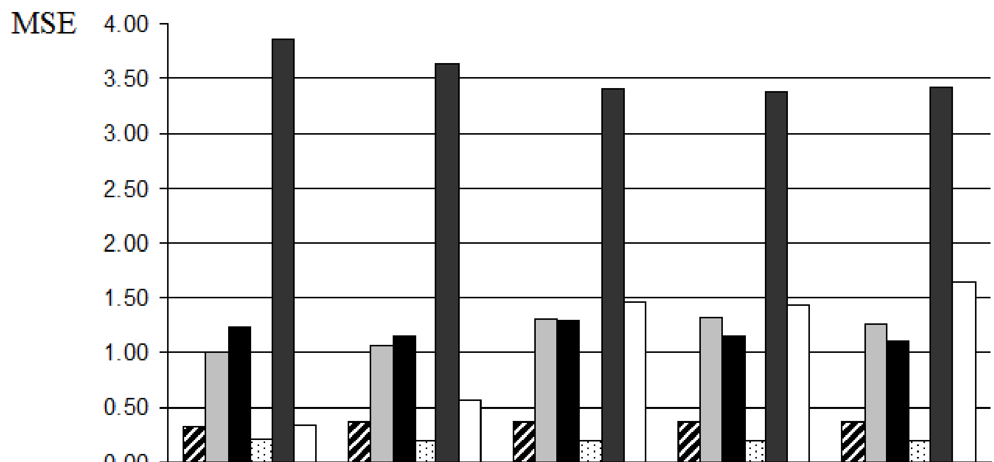
We first report on the sensitivity of our findings to changes in the foregone payoff calculation given by equation (1) which is used by the evolutionary operators of the IEL. Recall that for the IEL simulations reported in section 3.4, the history consisted of the observed actions of others and recommendations for play from the two most recently played periods, i.e., $T = 2$. Here we consider increasing the history length used for the evaluation of foregone payoffs from $T = 2$ to $T = 4$, 6, 8 or 10, which is implemented by changing the upper bound to the summations in equation (1). The idea is that a longer history length, T, may help to promote learning. However, the way in which the longer history length is evaluated may be temporarily destabilizing with regard to evaluation of foregone payoffs, as the number of periods over which foregone payoffs are evaluated increases steadily until the upper bound of T periods is reached. In addition, the random initialization remains a source of noisy initial evaluation of foregone payoffs. However, once strategies can be repeatedly evaluated in a stationary history of length T, then the algorithm is able to coordinate rather quickly.

The mean squared errors (MSEs) between the simulated payoffs and the human subject payoffs for the Battle of the Sexes game for these different history lengths are shown in Figure 3.8; analogous results for the Chicken game are shown in Figure 3.9.

For the Battle of the Sexes game (Figure 3.8) there is little or no improvement in MSE from allowing for more than 2 periods of evaluation in most treatments. For the Strangers Indirect Recommendation treatment alone, a longer history length works to *increase* the MSE, because this longer history allows agents to learn to ignore recommendations quicker and coordinate in a manner that is similar to the None Random treatment (although slower and less efficiently since the probability of following recommendations does not fall all the way to zero).

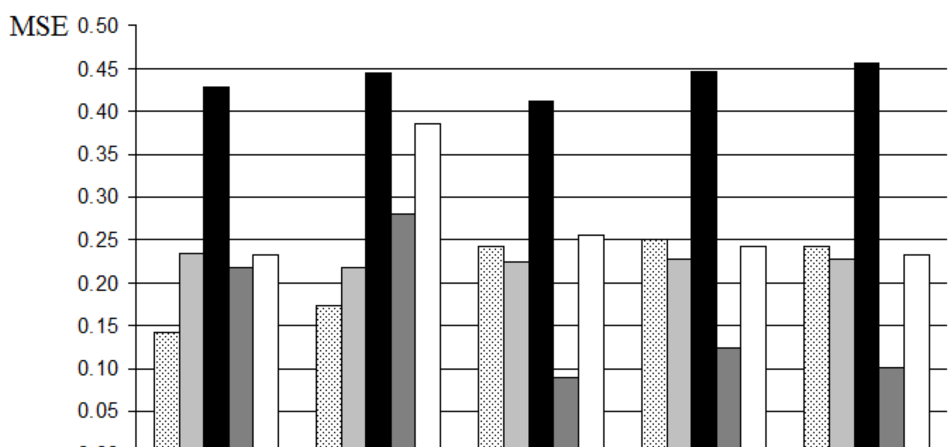
As for Chicken, increasing the length of information used to evaluate strategies also has little or no effect on MSE, with the exception of the Very Good recommendations treatment. In that treatment, longer histories help agents to learn to avoid following recommendations (as they should); they learn to follow them with too high a frequency in the baseline $T = 2$

[12] We have considered alternatives measures for assessing fitness such as the period-by-period frequency of play of actions C or D, which gives similar results to period-by-period payoffs. Fitting to the frequency of recommendation following does not perform as well as fitting based on payoff information, in part because initial conditions for recommendation following in the experimental data can differ from IEL's random initialization to a large degree, but also because not all treatments have recommendations.



	T = 2	T = 4	T = 6	T = 8	T = 10
Partners DR	0.3341	0.3671	0.3619	0.3675	0.3731
Partners None	1.0076	1.0590	1.3082	1.3243	1.2622
Partners IR	1.2355	1.1488	1.2834	1.1442	1.1084
Strangers DR	0.2118	0.1943	0.1924	0.1960	0.1952
Strangers None	3.8669	3.6310	3.4083	3.3808	3.4279
Strangers IR	0.3356	0.5599	1.4601	1.4367	1.6417

Figure 3.8: MSE Between Simulations and Experimental Data for Different Evaluation Lengths, Battle of the Sexes Treatments



	T = 2	T = 4	T = 6	T = 8	T = 10
Good	0.1418	0.1726	0.2429	0.2505	0.2411
Nash	0.2340	0.2170	0.2244	0.2269	0.2275
Bad	0.4269	0.4441	0.4117	0.4453	0.4562
V. Good	0.2174	0.2811	0.0884	0.1228	0.1011
None	0.2320	0.3851	0.2559	0.2424	0.2320

Figure 3.9: MSE Between Simulations and Experimental Data for Different Evaluation Lengths, Chicken Treatments

specification. Notice further that the MSE for the Chicken $T = 2$ treatments are small in magnitude as compared with BoS (the vertical scales are quite different).[13] Thus, for the game of Chicken, further reductions in MSE as T is increased are harder to come by. We conclude that, with a couple of exceptions, our simulation results are not very sensitive to our restriction that foregone payoffs are evaluated using a history length of $T = 2$.

3.5.2 Initialization

Recall that the two parts of the J strategies of the IEL algorithm were randomly initialized as detailed in footnote 2. Random initialization is a reasonable assumption if one has no prior information on the distribution of play of the strategies available to players. In this section we consider how our simulation results would change if we had instead initialized strategies in such a way that they matched the payoffs earned by the human subjects in the initial periods of play. Our analysis is limited to the Chicken game because in the Battle of the Sexes game initial play in the no recommendations treatment is essentially uniformly random already. By contrast, for the Chicken game, initial play is not uniformly random. Thus, we are interested in whether the fit of IEL to the Chicken game treatments can be improved by initializing the set of strategies in such a way as to match the initial behavior of the human subjects. Specifically, since the distribution of play observed in the No recommendation treatment's initial periods of play had C chosen approximately 2/3 of the time and D chosen the remaining 1/3 of the time, we considered simulations where the Not Follow element of the J strategies is assigned a value of 1 (i.e., C) with probability 2/3 and a value of 0 (i.e., D) with probability 1/3.

We also considered adjusting the initial conditions for the Follow part of the Chicken game strategies. Specifically for the Follow part we employed a narrower distribution than the uniform random distribution over $[0, 1]$ used in the IEL simulations reported in the previous section. The distribution was narrowed so that the probability that a recommendation would be followed or not given the distribution of the Not Follow portion of the strategy approximates the initial frequency of recommendation following that was observed in the experimental setting. Specifically, the restrictions are based on the observed minimum and maximum frequency with which the group of experimental subjects followed the recommendation over the first four periods of each Chicken treatment with recommendations. Let μ be the probability an agent follows the recommendation, let ψ_{min} be the minimum observed frequency of rule following, let ψ_{max} be the maximum observed frequency of rule following, let Φ_C be the distribution of C in the Not Follow portion of the strategy with $\Phi_D = 1 - \Phi_C$ being the distribution of D in the Not Follow portion, and let r_C be the probability the recommendation is C with $r_D = 1 - r_C$ being the probability the recommendation is D. The observed frequency of recommendation following can be described as

[13]The lower MSEs in Chicken as compared with BoS are an artifact of the lower variance in payoffs in the Chicken game, as compared with Battle of the Sexes.

$r_C(\mu+(1-\mu)\Phi_C)+r_D(\mu+(1-\mu)\Phi_D)$. Taking into account ψ_{min} and ψ_{max}, the distribution of the Follow portion becomes:

$$f \sim U\Big[\frac{\psi_{min} - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}{1 - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}, \frac{\psi_{max} - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}{1 - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}\Big] \tag{3.3}$$

Thus the Follow Distribution is created to initialize the algorithm with a similar frequency of recommendation following as found in the initial periods of human subject play, given the distribution in the Not Follow portion.

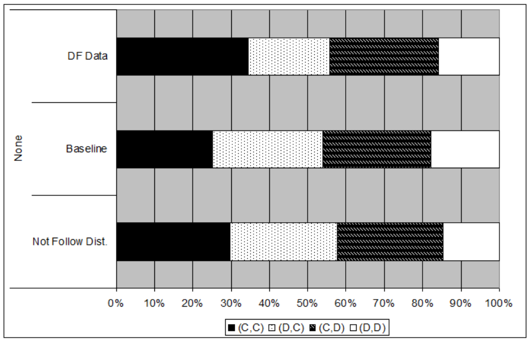
To better understand the marginal contribution of each initialization change, for each Chicken treatment, we conducted 500 runs of the IEL simulation with just the Not Follow part initialized to match the data (i.e., the Follow part remained a uniform random draw over [0,1]) or with just the Follow part initialized to match the data with r_C depending on the treatment (i.e., the Not Follow portion was initialized with each element being equally likely to play C or D, $\Phi_C = 0.5$) or where Both adjustments were simultaneously made to the initial Not Follow and Follow Distributions used in the simulations.

The results of these exercises are shown in Figure 3.10 which reports on the frequencies of outcomes in the DF experimental data and our simulations with various different initializations, including the Baseline, random initialization for the Not Follow and Follow portions of the strategies.

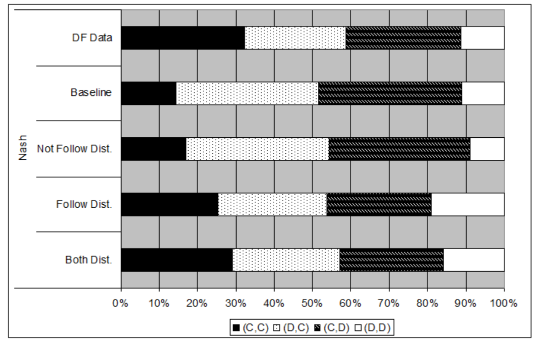
The figures reveal that for some, but not all treatments, a change in initial conditions for the Follow and Not Follow parts of the strategy so as to better match the initial experimental data, can yield improvements in the overall frequencies of outcomes so that these more closely resemble the human subjects data. For instance, in the Chicken-Nash treatment (panel (b) of Figure 3.10), the outcome frequencies from the IEL simulation provide a better match to the data if the initial conditions in the Not Follow or the Follow portions of the strategies are chosen so as to match the initial frequencies in the experimental data. Adding Both non-random initial conditions yields an even better fit. However, such improvements from non-random initial conditions do not always arise; for instance, in the Chicken-Good treatment, a departure from the baseline random initialization can result in a worse fit of the IEL simulated frequencies to the experimental data, e.g., when the initial conditions match the initial distribution of strategy choices in the Not Follow portion of the strategy.

3.5.3 Grid Search Description and Results

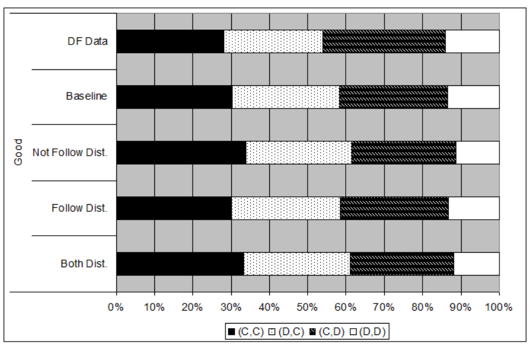
In this section we report on a grid search over the three main free parameters of the IEL with the aim of improving the fit of the model to the experimental data. We again use the mean squared errors of the period-by-period difference in payoffs between the simulated and experimental data to determine the goodness of fit, and we compare our findings with the "Baseline" parameterization used for the IEL simulation results reported in section 3.4,



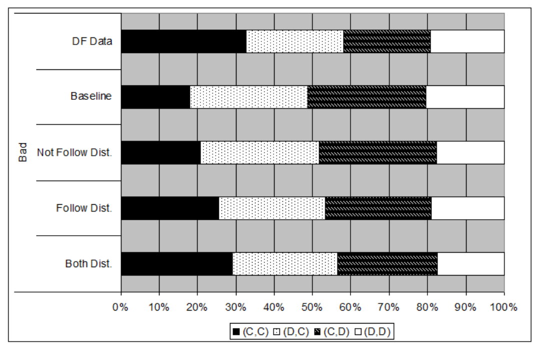
(a) Chicken None



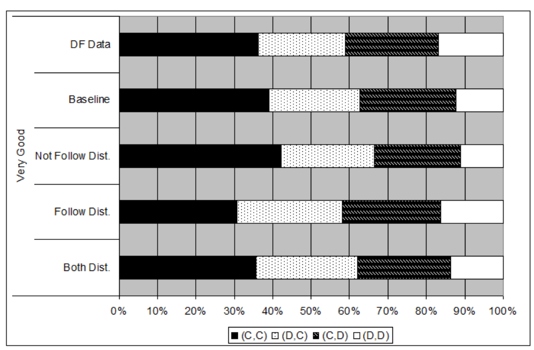
(b) Chicken Nash



(c) Chicken Good



(d) Chicken Bad



(e) Chicken Very Good

Figure 3.10: Simulation Results Relative to Experimental DF Data and Baseline Random Initialization: Not Follow Dist = Initial Strategy Choices in the Not Follow Portion set to Match Initial Data; Follow Dist = Initial Recommendation Following Frequency in the Follow Portion set to Match Initial Data; Both Dist = Both the Not Follow and Follow Portions set to Match Initial Data

$(J, \rho, \theta) = (180, 0.033, 0.10)$. The search is conducted for Chicken using the Nash recommendation data and for Battle of the Sexes using the Strangers Direct Recommendations data, as these two recommendations treatments are the most similar to one another. The free parameters found in the grid search for the Chicken-Nash and BoS Strangers DR treatments are then used to simulate the outcomes for all other treatments of, Chicken or BoS games, respectively. These simulated outcomes are reported in Appendix B for the interested reader. Here we merely report on the free parameter vectors obtained from our grid search, and the MSEs associated with using these best-fit parameter vectors in the other treatments in comparison with the baseline parameterization.

In our grid search, we first conduct a coarse grid search over the three parameters, J, ρ and θ. We then narrow the search around the values that minimize the mean squared error of the first search to determine if there is a further improvement to be made in the fit of the model to the data.

The grid searches for the Chicken and BoS games both implement the same 'coarse' search where J is varied from 150 to 250, in increments of 10, and both ρ and θ are varied from 0.05 to 0.95 in increments of 0.1. The coarse grid search results in the following best-fit IEL parameter values for Chicken: $(J, \rho, \theta) = (170, 0.250, 0.050)$, and for BoS: $(J, \rho, \theta) = (190, 0.050, 0.050)$. By contrast, the MSE for our Baseline parameter vector $(J, \rho, \theta) = (180, 0.033, 0.10)$ used for both games would have ranked 38^{th} in Chicken and 31^{st} in BoS out of the 1100 possible outcomes putting both within the top 5%.

We next conducted a finer grid search around the best-fit coarse grid search values. For Chicken, we varied J from 155 to 185 in increments of 5, we varied ρ from 0.15 to 0.35 in increments of 0.025, and we varied θ from 0.025 to 0.15 in increments of 0.025. This finer grid search did not yield an improvement over the coarse grid search, so the best-fit parameter values continued to be given by $(J, \rho, \theta) = (170, 0.250, 0.050)$ with an MSE of 0.2127. For BoS, we varied J from 175 to 205 in increments of 5, and we varied ρ and θ from 0.025 to 0.15 in increments of 0.025. This finer search resulted in a slightly different best-fit parameter vector (relative to the coarse grid search) of $(J, \rho, \theta) = (190, 0.050, 0.125)$ with an MSE of 0.1706.

	Good	Nash	Bad	Very Good	None
Baseline	0.1418	0.2340	0.4269	0.2174	0.2320
Grid Search Best-Fit	0.1716	0.2127	0.4241	0.2832	0.2339

Table 3.12: Comparison MSE Chicken

Tables 3.12 and 3.13 show the MSE for each of the treatments when using the baseline parameterization versus the best fitting parameter values from the finer grid search. Under the best-fit parameterization for Chicken (based on the Chicken-Nash treatment), both the Nash and the Bad Chicken treatments see an improvement in their MSE relative to the baseline MSE, while for the other three Chicken treatments, the MSE increases relative to the baseline MSE. Using the best-fit parameter vector for the BoS game (based on the Strangers Direct Recommendations treatment), the Strangers Direct Recommendations treatment as well as both of the None treatments have an improved MSE relative to the baseline parameterization whereas the Partners Direct Recommendation and both Indirect Recommendation treatments see an increases in MSE relative to the Baseline value.

	Partners			Strangers		
	DR	None	IR	DR	None	IR
Baseline	0.3341	1.0076	1.2355	0.2118	3.8669	0.3356
Grid Search Best-Fit	0.3498	0.9198	1.5789	0.1706	3.7083	0.4517

Table 3.13: Comparison MSE BoS

These results suggest that while it is possible to find IEL parameterizations that improve upon the one we used for the IEL simulations reported on in Section 3.4, the improvement in MSE for the targeted treatment does not necessarily carry over to other treatments of the same game using that same optimized parameter vector. We note further that our baseline parameterization is not so far off from the best-fit parameter vectors, with the possible exception of the ρ parameter for the Chicken game simulations where the best-fitting value of 0.25 is higher than our baseline choice of 0.033.[14] The higher value for ρ in the Chicken game appears due to the noise generated by the human subjects' difficulty in coordinating on an equilibrium in the Chicken-Nash treatment.

3.5.4 Sensitivity of Simulation Results to Single Parameter Changes

Our final sensitivity exercise examines how the fit of the IEL model to the data changes as we perturb a single free parameter of the model away from the baseline parameter vector $(J, \rho, \theta) = (180, 0.033, 0.10)$ used in the simulations reported on in section 3.4. Note this is a different exercise than the grid search, which varied all three parameter values at once in a search for the best fitting parameter vector. Table 3.14 reports on the effects of adjusting a single IEL model parameter below or above the Baseline parameter model choices on the period-by-period MSE between the simulated and actual payoffs. As in the grid search exercise of the previous section, we focus on the BoS Strangers Direct Recommendations treatment and the Chicken Nash Recommendations treatments for this exercise. For both of these treatments, using fewer strategies, e.g., $J = 100$ or a smaller ρ, while holding the other two IEL model parameters at baseline values would improve the fit of the IEL model to the data as indicated by the lower MSEs relative to the Baseline but the differences are small.

[14]The 2^{nd}, 3^{rd}, and 4^{th} best outcomes of the coarse grid search all had $J = 190$ and $\rho = 0.05$ with θ being 0.25, 0.35, and 0.15, respectively. We ran an additional grid search using the second best outcome of $(J, \rho, \theta) = (190, 0.05, 0.25)$ where J varied from 155 to 185 in increments of 5, ρ varied from 0.025 to 0.150 in increments of 0.025, and θ varied from 0.150 to 0.350 in increments of 0.025. The outcome of this search revealed that $(J, \rho, \theta) = (170, 0.025, 0.175)$ produced a smaller MSE = 0.2091. The latter parameterization creates the same trade-off of improving some, but not all outcomes for the Chicken game.

For BoS a larger set of strategies $J = 250$ would provide an even greater improvement in the fist of the IEL model to the data, so the implications for the set of strategies is inconclusive for the BoS game. The largest consistent improvement comes from using a smaller ρ which results in an $\approx$ 8.5% decrease in the MSE for BoS and an $\approx$ 1.6% decrease in the MSE for Chicken. All of the other adjustments reported on in Table 3.14 would come with the tradeoff of improving the fit of the model in one game while reducing the fit in the other game.

	BoS Strangers Direct Recommendations						
	Baseline	$J = 100$	$J = 250$	$\rho = 0.01$	$\rho = 0.10$	$\theta = 0.01$	$\theta = 0.20$
MSE	0.2118	0.2009	0.1886	0.1937	0.2135	0.2104	0.2139
	Chicken Nash Recommendations						
	Baseline	$J = 100$	$J = 250$	$\rho = 0.01$	$\rho = 0.10$	$\theta = 0.01$	$\theta = 0.20$
MSE	0.2340	0.2302	0.2398	0.2303	0.2328	0.2429	0.2290

Table 3.14: Effects of Single Parameter Changes on MSE

We conclude from this exercise that there is not much consistency in the effect of varying a single IEL parameter to improve the fit of the model to the data across the two games, and that the improvements in MSE relative to the baseline MSE are not so large that we are missing much from simulations conducted using the baseline parameterization.

3.6 Conclusions

The learning in games literature has mainly focused on the conditions under which Nash equilibrium can be learned. By contrast, there has been little research exploring whether and how *correlated equilibria* can be learned which is surprising since the correlated equilibrium concept is an important generalization of the Nash equilibrium concept and correlated equilibria can be more efficient than Nash equilibria. From a policy perspective, the optimal system design may be characterized by a correlated equilibrium as opposed to a Nash equilibrium; the traffic light signal is the canonical example, but there are other examples as well.[15]

In this chapter, we have shown how an individual evolutionary learning algorithm can be adapted to learn a correlated equilibrium in a finite amount of time. Importantly, our algorithm does not require knowledge of the correlated distribution of recommended actions, ex-ante. We believe that this result is new to the learning in games literature. Our behavioral algorithm, based on evolutionary concepts, often learns to quickly rely upon a recommendation (or signal) whereas the theoretical literature demonstrates learning only in the limit and only to a set of correlated equilibria. In comparisons between our simulations

[15]In Chicken for example, a property rights policy wherein the first player to arrive at a location (the incumbent) plays D and the late arriving party (challenger) plays C is one that corresponds to the Nash recommendations correlated equilibrium, see Gintis (2009).

and the experimental data, we find that the IEL algorithm is able to mimic the behavior of human subjects in many, though not all experimental treatments, providing some external validity for our approach. A sensitivity analysis shows that our findings are largely robust to alternative specifications and parameterizations of the model. The IEL algorithm is not able to effectively interpret indirect signals as the human subjects were able to in the Battle of the Sexes game, and IEL does not condition on differences in matching protocols (Partners versus Strangers) which were known to the human subjects and affected their behavior. Another difference, as mentioned in Bone et al. (2013) is that experimental subjects might be concerned with the payoffs earned by other subjects. Such other regarding concerns may provide an alternative explanation as to why the human subjects in the Chicken experiment are sometimes less willing to follow recommendations than our simulated agents. We leave further examination of these issues to future research.

Chapter 4

Individual Evolutionary Learning with Beliefs and Other-Regarding Preferences

4.1 Introduction

Social dilemmas represent an interest for economist, especially experimental economists, because a division exists between individually self-interested behavior and socially optimal choices. The Linear Public Goods Game provides a relatively simple method to examine whether agents adopt the self-interested strategy or play some other strategy when facing this sort of dilemma. A sizable proportion of the experimental participants involved in these games adopt strategies that are not consistent with either the individually self-interested strategy or social optimal strategy. While there are many theories that can possibly explain this behavior, this chapter seeks to extend the Individual Evolutionary Learning with Other-Regarding Preferences (IELORP*) by Arifovic & Ledyard (2012) through the introduction of beliefs and mimicking the experimental designs more closely. The IELORP* has been very successful at fitting the data while remaining consistent with the many set of stylized facts associated with the Linear Public Goods Game.

I show how extending the model by introducing beliefs based on adaptive expectations and implementing the discrete choices used in the experimental games continues to match the data at levels consistent with the previous version of the IELORP*. The weights implemented for the adaptive expectations are determined using the data of Gächter & Renner (2010) via a simulation method and an OLS regression. While both methods produce similar contribution level data, the weights from the OLS regression provide a better fit to the belief data. This is likely due to fact that the OLS captures the heterogeneity of the individuals in the experiment since the OLS model relies on fixed effects whereas the simulation method ignores heterogeneity. By introducing beliefs with a greater level of sophistication than purely myopic beliefs, the model can be applied to games where beliefs are elicited via a monetary reward with relatively minor changes to the algorithm. The model captures the general behaviors observed in the Linear Public Goods Game, but not to the same degree as when belief elicitation is not monetized. The lack of fit appears to be partially related to the experimental setting and design that is not taken into account using the adaptive expectations. The influence of the experimental design is illustrated using a version of the IELORP* that treats stated beliefs and contribution choices as entirely separate elements in each agent's set of strategies.

4.1.1 Related Literature

The related literature section is divided into two main components: a) the literature on social dilemmas, particularly experimental analysis and application of the Linear Public Goods Games; and, b) the literature on agent-based modeling.

Social Dilemmas

Table 4.1: Prisoner's Dilemma

		Column Player	
		Cooperate	Defect
Row	Cooperate	0,0	-4,1
Player	Defect	1,-4	-2,-2

Robert Axelrod set out a grand goal for the social sciences in his pioneering book *The Evolution of Cooperation*. The book begins by asking "[u]nder what conditions will cooperation emerge in a world of egoists without central planning?" (Axelrod (1984), p.3). To illustrate the issue of cooperation, Axelrod focuses on the Prisoner's Dilemma to introduce the concept of a social dilemma. The Prisoner's Dilemma shown in Table 4.1 provides an illustration of the issues inherent in all social dilemmas. Social Dilemmas are interactions amongst multiple people generally modeled as a game where the rational strategy to adopt is determined by using either an individual or a social perspective. The individually self-interested behavior here is to use the Nash Equilibrium since playing "defect" guarantees the highest payoff achievable for the individual regardless of the other player's action (i.e. $1 > 0$ and $-2 > -4$). Whereas the socially optimal strategy from a Social Planner's perspective is to "cooperate" with your opponent providing the highest aggregate payoff (i.e. $1 + 1 = 2 > -3$ or -4). The results from experiments show that individuals playing the Prisoner's Dilemma are far more likely to play "cooperate" than would be theoretically predicted based on individual self-interest. Experimental economist began to allow the subjects to play the game repeatedly to further their understanding of individual behavior, but the subjects continued to play "cooperate" more frequently than Nash Equilibrium predicts especially in the early periods of play. The Prisoner's Dilemma is only able to provide a limited view of cooperation as the focus is on the behavior of two individuals interacting with each other when their choices are restricted to two binary decisions.

The Linear Public Goods Game provides a more robust example of a social dilemma to examine the formation of cooperative behavior. The Linear Public Goods Game has agents select a proportion of an endowment to contribute to the public good based on the following payoff function:

$$\pi^i = w^i - c^i + M \sum_{j=1}^{N} c^j \ \forall i = 1, ..., N \tag{4.1}$$

where w^i is the initial endowment of each agent, $c^i \in [0, w^i]$ is the choice of contribution made by the agent, and M is the constant rate of return from contributing to the public good with $(1/N) < M < 1$. The Linear Public Goods Game normally includes more than two players and allows for a number of alternative choices by the players (i.e. how to distribute some initial endowment)[1]. Despite these differences with the Prisoner's Dilemma, the Linear Public Goods Game has the exact same predicted results. In other words, the individually self-interested strategy is to contribute nothing (equivalent to "defect") and the socially optimal outcome is for all players to contribute the endowment w^i (equivalent to "co-operate"). Much like the experimental outcomes of the Prisoner's Dilemma, many individuals contribute some non-zero proportion of their endowment to the game instead of behaving in a manner consistent with the Nash Equilibrium prediction. The reason for this outcome is still debated with a contingent of experimental economists insisting the behavior being captured can all be attributed to confusion by the experimental subjects[2]. To develop a better understanding of how participants play (and determine if confusion) is the reason the Nash Equilibrium prediction was not used by all of the experimental participants, experimental economists have moved towards a repeated version of the game such that $\Pi^i = \sum_{t=1}^{T} \pi_t^i$ where Π^i describes an individual's total payoff from playing the experiment for T periods. A sizable proportion of the experimental participants continued to exhibit behavior inconsistent with predictions of the Nash Equilibrium, while the average contribution observed over time does move closer to the Nash Equilibrium prediction. The change in behavior over time prompted suggestions that the behavior might not be the result of confusion[3], but represents some other form of behavior such as learning or conditional cooperation on the part of the experimental participants. Due to the many experiments conducted using the Linear Public Goods Game (hereafter LPGG), a number of survey

[1] This is not always the case, for example Wilcox & Feltovich (2000) have a Linear Public Goods Game where individuals make a binary decision over whether to contribute their endowment or not.

[2] See Burton-Chellew et al (2016) for a recent example of the confusion argument.

[3] See Bayer et al (2013) or Keser (1996) for experiments providing evidence contradicting the confusion argument for the repeated version of the Linear Public Good Game.

articles have been written on the subject (See Ledyard (1995), Holt & Laury (2002), and Chaudhuri (2011)[4]).

Based on the observations from the many Linear Public Goods Games experiments, a consistent set of stylized facts has been associated with the Linear Public Goods Game (See Arifovic & Ledyard (2012) or Klumpp (2012)). These stylized facts are that average contribution begins at about 50% then declines with repetition, there does not appear to be a monotonic pattern of contribution by individuals over time, increasing the value of M leads to increases in the average contributions (particularly for small N), increasing the groups size leads to increases in the average contribution (particularly for small M), and an unanticipated "restart" of the experiment after time T leads to an increase in average contribution at $T + 1$. Any theory of behavior in the LPGG must be able to adequately address these areas.

More recently, experimental economists have moved beyond simply asking experimental subjects to make contribution choices and have begun trying to elicit the underlying beliefs of the experimental subjects in the hopes of understanding the motivation of experimental subjects. Understanding the influence and importance of beliefs has been a general interest of economist due to their importance in economic theory since before Savage proposed the Subjective Expected Utility showing a formal importance of beliefs. Several recent experiments have elicited the beliefs of agents while playing the LPGG (See Croson (2000), Gächter & Renner (2010), Fischbacher & Gächter (2010), Neugebauer et al (2009), Chaudhuri & Paichayontvijit (2010), and Smith (2013))[5], yet questions remain as to how the belief elicitation affects play in an experimental settings when monetary incentives are utilized. While the number of papers examining this area has been growing, there still does not appear to be sufficient evidence to determine if the behavior of individuals participating in experiments changes when individuals are paid to express their beliefs. For instance, Ruström & Wilcox (2009) find support for incentivized elicitation of beliefs changing the behavior of individuals in the context of a Matching Pennies Game while Nyarko & Schotter (2002) do not. Within the context of the Linear Public Goods Game there also exists contradictory evidence with Croson (2000) finding a lower average contribution to the public good (i.e. closer to "defection") with incentivized belief elicitation and Gächter & Renner (2010) finding an increase in the average contribution to the public good (i.e. closer to "co-operate") with incentivized belief elicitation. While Wilcox & Feltovich (2000) find no support for a change in behavior of incentivized elicited beliefs, their incentive mechanism was not based on a proper scoring rule and only allowed for a binary contribution choice.

[4]Several other important studies should be mentioned, for instance Andreoni & Croson (2008) review the literature on "Partners" vs. "Strangers" amongst other things, Fehr & Gächter (2000) give an overview of punishment within this general framework, and Zelmer (2003) gives a meta-analysis of the public goods experiments, which is suggestive as to the amount of data/experimental analysis that has taken place.

[5]Schlag et al (2015) provides a relatively comprehensive survey of belief elicitation including the Linear Public Goods Game.

While not yet considered to be among the stylized facts, all of the experiments examined find *overly optimistic* beliefs discussed by Neugebauer et al (2009). The beliefs are overly optimistic because the average stated belief is higher than the average contribution choice over the whole experiment. Note that this does not mean that the stated beliefs are overly optimistic each period.

Simulated Behavior in the Linear Public Goods Game

Simulations have become a useful method to illustrate, explain, and understand more complex behavior in the Linear Public Goods Game. Clemens & Riechmann (2006) implement an evolutionary algorithm to explain the lack of full free-riding behavior in LPGG experiments, but does not fit any specific set of experimental data. In addition to an experiment, Fischbacher & Gächter (2010) created a set of simulated agents using two different elements from their experiment. The data generated from the experimental subjects of Fischbacher & Gächter was a necessary element in the simulation model, this limits the ability of the simulations to have any external validity. Wendel & Oppenheimer (2010) showed a set of conditions where simulations are able to generate "saw tooth" behavior seen in some experiments. Wendel & Oppenheimer suggested the experimental behavior is an attempt to signal other agents their co-operative nature, but the signal was noisy and thus not well understood. One important element from the simulations of Wendel & Oppenheimer was the necessity of heterogeneity to generate the "saw tooth" behavior. Arifovic & Ledyard (2012) presented a learning algorithm that produces simulated outcomes consistent with a number of stylized facts found in repeated LPGG with symmetric endowments, $w^i = w \, \forall i$. The results of simulations approximated the experimental outcomes discussed in Isaac & Walker (1988), Andreoni (1995), Isaac et al (1994), Croson (1996), and Andreoni (1988) utilizing a single set of parameter values. One appealing aspect of the approach taken by Arifovic & Ledyard is the attempt to create a model that is transferable to other experimental games while maintaining a consistent set of parameter values (i.e. externally valid). Bayer et al (2013) conducted an experiment to examine whether confusion could be eliminated as the explanation for behavior in the LPGG. To examine the difference in behavior Bayer et al rely on simulations using reinforcement learning to differentiate simple learning behavior from confusion. Cotla & Petrie (2015) focused on the social preferences aspects of the LPGG and utilized simulations to suggest that some papers over-emphasize social preferences while other under-emphasize them. Cotla & Petrie found that reinforcement learning with the initial behavior described by social preferences was best able to match the data from their experiment.

Given the transferability of the Arifovic & Ledyard model and its ability to fit the stylized facts of the LPGG, this chapter will focus on extending the Individual Evolutionary Learning Model with Other Regarding Preferences. A brief overview of the model is provided here with a more detailed description given later in this chapter. The first part of IELORP* is the idea that agents enter the experiment with Other Regarding Preferences (ORP). This implies that payoffs previously described by π^i do not actually describe the payoffs experienced by the agents. The assumed utility function is:

$$u^i(c) = \pi^i(c) + \beta^i \overline{\pi}^i - \gamma^i \max\{0, \overline{\pi}(c) - \pi^i(c)\} \tag{4.2}$$

The $\pi^i(c)$ is the individual payoff as previously described, $\overline{\pi}(c)$ is the average payoff of the group, β^i is the level of altruism by the agent, and γ^i is the (one-sided) envy an agent feels when being "taken advantage of." Further there is the assumption that agents are heterogeneous in their levels of β and γ. This produces three Nash Equilibria strategies (contribute nothing, everything, or the average) that depend on the number of agents, the marginal per capita return, each agent's level of altruism, and each agent's level of envy. Thus for all agents if $\frac{1-M}{M-\frac{1}{N}} \leq 0$ the agent will behave as the individually self-interested theory suggests and free ride on the contributions of other agents. The difference between an agent who contributes everything or one who acts as a conditional co-operator will be determined by whether $\gamma^i(\frac{N-1}{N}) \geq (\text{or} \leq)[(M - \frac{1}{N})\beta^i + M - 1]$ (respectively). These results are for the equilibria of the stage game and do not take into account the learning aspects of the model.

The second part of IELORP* is the existence of behavioral model called Individual Evolutionary Learning (IEL) based on a genetic algorithm. Given the utility specification, the LPGG can be modeled as a stage game G played for T periods as done by Arifovic & Ledyard (2012). The stage game can be described by $G = N, X, \pi, r$ where N is the number of players, X represents the action space, π describes the payoff function based on the joint strategies chosen, and r gives the information that is available to the agent. Given this game, the agent's behavior can be described by (A^i_t, ψ^i_t) where A^i_t is the set of strategies agent i has available at time t and ψ^i_t is a probability measure on A^i_t describing the likelihood any strategy will be played. After selecting a strategy from the set A^i_t and getting feedback in terms of π^i_t and $r^i_t = \sum_{\ell \neq i} c^\ell_t = \mu^i_t$ the agent will then update their strategy set based the experimentation, replication, and selection genetic operators. In addition to these elements, there is an initialization of the agents A^i_t and method to determine the initial strategy used[6].

Finally, IELORP* puts both the ORP and IEL together to create a behavior model that replicates the five stylized facts from the experimental literature. Given the behavioral assumptions about the game, this means that the functional form for utility can be modified to:

$$\begin{aligned} u^i(c^i|\mu^i) = {} & [(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{i*}(\frac{N-1}{N})]c^i \\ & + [M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^{i*}}{N}](N-1)\mu^i + (1+\beta^i)w \end{aligned} \tag{4.3}$$

[6]More description of the genetic operators and the implementation of the algorithm will take place when describing how beliefs are implemented in the model itself.

where μ^i is the information on the average contribution of others and $\gamma^{i*} = \gamma^i$ if $\overline{\pi} \geq \pi^i$ or 0 otherwise. This describes how the information is incorporated, but does not complete the description of the model. A complete description requires discussing several elements that will be left for later when the details of the Individual Evolutionary Learning model are given.

The heterogeneity of ORP exists in the model as the triple (P, B, G) where P represents the proportion of the population that is purely selfish, B represents the max of a uniform distribution for altruism over $[0, B]$ and G represents the max of a uniform distribution for envy over $[0, G]$. The choice of (P, B, G) that provided the best fit for the Isaac & Walker (1988) data was $(0.48, 22, 8)$. The same choice of (P, B, G) was also used in the simulations showing the transferability of the model and are adopted in this chapter as well.

4.1.2 Contribution of this Chapter

This chapter seeks to introduce a more sophisticated version of the IELORP* by matching the experimental setting as closely as possible while also introducing beliefs into the model. By adjusting the IELORP* to account for the discrete nature of the experiments and moving beyond static beliefs, the goal is to demonstrate these changes continue to capture the behavior seen in the experimental settings while allowing the model to be extended to investigate more recent experiments. The experiments have their subjects make choices over discrete values whereas the IELORP* allows the simulated agents to use any value over $[0, w]$. As algorithmic approaches can get caught at local extremes moving to discrete values could have resulted in simulated agents becoming stuck on a particular contribution choice regardless of the input from other players, but this is not the case here. The beliefs take the form of adaptive expectations with the weight on the model estimated from the data of Gächter & Renner (2010).

The introduction of beliefs also allows for a more sophisticated initialization of the model than previously possible. The IELORP* simply selects a contribution choice at random with probability $1/J$. Since the initial set of contributions used J draws from the uniform distribution over $[0, w]$, the overall contribution choice will be approximately 50%. This outcome matches nicely with the data, but not perfectly as experiments relying on a lower value of M generally have outcomes lower than experiments with a higher M (ignoring the influence of N). This can be clearly seen in the Isaac & Walker (1988) results. Cotla & Petrie (2015) found their pay-off based reinforcement learning model was able to match their experimental data when initialized via social preferences. Arifovic et al (2016) also showed how adjusting the initialization can vastly improve the fit of simulations to experimental data. While both Cotla & Petrie and Arifovic et al rely on the experimental data to generate the adjustments to the initialization, the model in this chapter assumes agents utilize a (false) consensus to determine their initial behavior. This modification of beliefs does appear to improve the outcome some, but does underestimate the influence of M seen in the experimental results and completely ignores the number of agents N.

To test if the changes made to the algorithm break the model, the outcomes are compared to results reported by Arifovic & Ledyard (2012). The chapter demonstrates that the influence of using discrete choices, adjusting the initialization, and introducing beliefs is minimal when the beliefs are not elicited via monetary rewards. This means that the adjustments do not destroy the transferability of the model. Because my model includes beliefs, the model can be compared to the experiments utilizing belief elicitation. When monetary rewards are introduced via proper scoring rules, the model requires some minor adjustments to account for the additional payoff. Even with these adjustments, the adaptive beliefs are unable to fit the data as well as before suggesting the transferability may have broken down. The model is able to match the experimental results of Gächter & Renner, but not the results of Croson (2000) or Neugebauer et al (2009). To give a point of comparison, an alternative algorithm is introduced that allows for a greater freedom between stated beliefs and contribution choices. This model better accounts for elements in the experimental design, but is still unable to match the data very well.

4.2 IELORP* and "Un-"elicited Beliefs

While the IELORP* does an excellent job of matching the aggregate behavior seen in many experiments (see Arifovic & Ledyard (2007) or Arifovic et al (2015) for examples), it relies on behavior from the previous period to determine what action to adopt in the next period (i.e. the agents are myopic). For this reason the initialization process in the IELORP* involves each element of A_t^i being drawn with uniform probability from a continuous distribution over $[0, w]$ (as opposed to being limited to the integer values as was done in the experimental settings). Once the set of strategies has been created, the choice of which strategy to use is determined by giving an equal probability to each strategy within A_t^i. In addition, all future choices are entirely dependent on what the average outcome was from the previous period, i.e. the agents have myopic beliefs and expect the play of other agents in the next period to match what happened this period. The goal of this section is to introduce an underlying beliefs system, un-reflected beliefs[7], about the future while continuing to allow the IELORP* to match the five stylized facts from the experimental data. The beliefs will be modeled via adaptive expectations. After describing the form beliefs will take in this chapter, the details of the algorithms will be discussed including how beliefs are created and how the algorithm is initiated. Finally, this section will conclude by comparing the results of the algorithm with experimental data and the previous IELORP* model.

[7]The view of beliefs not being questioned unless there is a significant reason to do so is consistent with psychological studies as well e.g. Kahnemen (2011)

4.2.1 Un-reflected Belief Learning

As will be discussed in the third section, papers like Croson (2000), Ruström & Wilcox (2009), and Gächter & Renner (2010)[8] elicited beliefs via a proper scoring rule and examined how this may change participants' behavior in an experimental setting. Before considering this issue, there exists a question of what the agents are doing or believing when there is no elicitation of beliefs. While beliefs elicited via incentivized mechanisms are clearly part of the strategy set of an agent, the focus of this section will be on examining behavior when no incentive mechanism exists. The suggestion in this chapter is that un-incentivized beliefs are not strategic. Although the beliefs an agent has will clearly affect their choice of strategy, this insight simply implies that beliefs are generally an unconscious element in an agent's process of learning. Assuming that agents are indeed not fully aware of where or how their beliefs are formed when not incentivized, the idea implemented here is that beliefs will follow a routine when being updated.

The idea that the beliefs are updated unconsciously, however, does not specify a particular form the beliefs should take. The only implication is that they may follow some particular trend and the exact form this trend takes is unclear. Ruström & Wilcox (2009) found that those agents who did not have their beliefs elicited eventually start playing like those who did have their beliefs incentivized. Based on this, the beliefs of incentivized agents and the beliefs of agents who have not explicitly stated their beliefs will converge given sufficient time. The overall implication is that the agents with un-elicited beliefs are formed unconsciously making learning much slower. While there is no data supporting this idea for the LPGG, the overall idea that the beliefs when not elicited are formed unconsciously will serve as a basic assumption of this model. I will take the view that the beliefs can be modeled as a version of adaptive expectations with $\mathrm{E}[b^i_{t+1}] = \mathrm{E}[b^i_t] + \lambda * (\mu^i_t - \mathrm{E}[b^i_t])$ where $\mathrm{E}[b^i_{t+1}]$ is the expectation or belief of the average contribution of other agents next period, $\mathrm{E}[b^i_t]$ is the currently held belief about average contribution of others, and μ^i_t is the actual (observed) average contribution by others. Note the similarity to the idea of "Natural Expectations" as modeled by Fuster, Laibson, & Mendel (2010) where agents have both *intuitive expectations* and *rational expectations.* I will assume the reader is sufficiently familiar with rational expectations. The intuitive expectations on the other hand are intended to take into account the idea of both an *anchoring bias* and *availability bias* discussed in psychological literature. In this model, an agent's current beliefs $\mathrm{E}[b^i_t]$ suffer from an anchoring bias[9], whereas the ex-post rational belief is μ^i_t. The two forms of expectation are weighted to create the *natural expectations* that an agent would have. More importantly, the authors make the case that this form of expectations is an "as if" process that is not consciously determined. Their paper takes that idea that λ represents an index of imperfection with

[8]See also Gneezy, Meier, & Rey-Biel (2011) for a more general view on this.

[9]The initial belief $\mathrm{E}[b^i_0]$ is the source of the achoring bias because the influence only fully disappears at the limit.

$\lambda \in [0, 1]$. The suggestion of natural expectations can then be viewed as a belief updating routine:

$$\mathrm{E}[b^i_{t+1}] = (1 - \lambda) * \mathrm{E}[b^i_t] + \lambda * \mu^i_t \tag{4.4}$$

In other words, the information that agents use is based on their belief of what other agents will do, i.e. $r^i(c_t) = b^i_t$.[10]

The value of λ is estimated using two different techniques applied to the experimental data from Gächter & Renner (2010). First, I use simulation to find a λ that minimizes the mean squared error between the actual $\mathrm{E}[b^i_{t+1}]$ and the estimated b^i_{t+1} from the right hand side of Eqn. (4.4), mimicking the methodology used in Branch (2004). The simulation consists of two grid searches, a rough grid search from zero to one in increments of 0.1 and a fine grid search from 0.4 to 0.6 in increments of 0.005[11]. This resulted in a $\lambda = 0.545$. Second, I estimate λ using OLS from the following equation: $(\mathrm{E}[b^i_{t+1}] - \mathrm{E}[b^i_t]) = \lambda * (\mu^i_t - \mathrm{E}[b^i_t]) + \delta^i + e$, where δ^i is an individual fixed effect. The choice of model selected here ensures the sum of the estimated coefficients totals one, whereas estimating via Eqn. (4.4) would allow for an outcome that does not sum to one[12]. While this method ignores the simultaneity issues discussed in Smith (2013, 2015), there is no causal argument being presented here since the assumption is purely instrumental. The result is $\lambda = 0.306$[13]. Because it is not clear which method should be preferred, the algorithm is run using both[14].

An important issue remains to be considered beyond the form of belief updating. Since the IEL learning has agents with J possible strategies, the obvious question is: Is there a single belief for each agent or are the J beliefs? This chapter is assuming that belief formation is an internal process that is unexamined, but this assumption by itself is not enough to answer the question. While a single belief makes evaluation of any expected payoff easier, it also restricts the performance of the agent by limiting their set of strategies to a single idea. This restriction runs contrary to the trial and error learning based on evaluation of different strategies. At the same time, having multiple beliefs about how others will behave complicates the calculation, evaluation, and comparison of expected payoffs since

[10] Additionally, Fuster, Laibson, & Mendel (2010) suggest that there is no differentiation between intuitive and naive expectations; here, however, the internal process used to create the initial anchor is markedly different from the new information being introduced into the belief system. Also, the traditional idea of naive or myopic expectations is consistent with the idea of static expectations, e.g. Muth (1985).

[11] Additionally, to allow for the possibility that heterogeneity might play a role, I conducted a search where λ was allowed to vary across agents. This resulted in a higher MSE than when λ was assumed to be a constant.

[12] When $1 - \lambda$ and λ are estimated using Eqn. (4.4), the outcome is close but not equal to one (nor statistical indistinguishable from one).

[13] The standard error was 0.058 making it statistically significant. This result is only reported to present the complete outcome of the OLS estimation and not intended to be used to draw inferences.

[14] The use of a single λ for all agents is a choice not taken lightly considering the IELORP* explicitly relies on heterogeneity in preferences; however, the 17 identical experiments run where beliefs were elicited without monetary rewards did not provide a sufficiently rich data set to develop a set of heterogeneous values for λ that provided a better fit of the Isaac & Walker (1988) data.

the expected payoffs could then vary over two potential elements of the strategy, i.e. replace μ^i with $b^i_{j,t+1}$ in Eqn. (4.3) and notice that the payoff from a strategy depends on both $c^i_{j,t+1}$ and $b^i_{j,t+1}$. The problem is illustrated when $c^i_{j,t+1} = c^i_{k,t+1}$ and $b^i_{j,t+1} > b^i_{k,t+1}$, then strategy j should provide a higher overall payoff based on the expected behavior of others. This chapter will assume that each agent has J strategies with each individual strategy consisting of both a belief $b^i_{j,t}$ and contribution choice $c^i_{j,t}$. Since the beliefs are assumed to be "unconscious," the average belief $\bar{b}^i_{t+1}$ is utilized to create the expected payoffs for the agent as the average captures the general belief held regarding the behavior of others.

4.2.2 Belief-based IELORP* Algorithm

In this section, I describe the elements of the learning algorithm based on the adaptive beliefs discussed in the previous subsection. The model uses the same evolutionary operators – experimentation, replication, and selection – as other IEL models. This model creates a set of expected payoffs to use during replication and selection rather than the hypothetical payoff function generally used in other Individual Evolutionary Learning models. The evolutionary process to create a new set of strategies is described with the steps of the algorithm being essentially unchanged from previous iterations. The initialization of the algorithm is altered and requires a thorough description.

The algorithm adheres to the following set of steps. First, the algorithm is initialized by creating a remembered strategy set A^i_1 for the first period. Next, a strategy is selected for play and the agent receives feedback on their payoff and the average payoff of others. The learning algorithm begins by updating beliefs and the evolutionary process starts with experimentation taking place to introduce new ideas into the set of strategies. The player then creates a set of expected payoffs based on their updated beliefs and contribution strategies. The expected payoffs are used to determine which strategies should remain in the remembered strategy set A^i_2 as well as used to determine which strategy will be implemented in the next period. The agent plays the selected strategy, receives feedback on their payoff and the average payoff of others, and abides by the learning algorithm for the remaining T periods. Thus the elements of the algorithm can be broken into two sections: 1) the initialization, and 2) the between period evolutionary process.

The Initialization

In each of the games, the agents make a discrete contribution choice without any knowledge of the behavior of the other players in the game. The remembered strategy set A^i_1 consists of J strategies for the choice of contribution and beliefs in the behavior of others. The contribution element of the strategy consists of J draws from the $[0, w]$ uniform distribution with the value drawn rounded to match the values available to the experimental subjects. The beliefs about the contributions of others are drawn from the uniform distribution from

$[0, (N-1)*w]$ and divided by $(N-1)$ then rounded to a value available to the experimental subjects[15]. This adjustment represents a departure from the continuous contribution choice over $[0, w]$ used in the original IELORP* model.

Unlike the IELORP*, the probability of playing each strategy ψ_1^i is not set to $1/J$. Since the agents have a set of beliefs about how others will behave, I assume that the agents are able to create a ψ_1^i based on the expected payoff from each strategy. This assumption is based on the idea of a (false) consensus effect where people are biased to view the behavior of others as similar to their own behavior[16]. Thus strategy j has a probability of being played equal to $\psi_{j,1}^i = \frac{u^i(c_j|\bar{b}_1^i)}{\sum_{k=1}^J u^i(c_k|\bar{b}_1^i)}$ with $\bar{b}^i$ replacing μ^i in Eqn. (4.3).

The Evolutionary Process

After a strategy has been played, the evolutionary portion of the algorithm begins. While the IEL style algorithm generally consists of 3 distinct elements, this chapter will layout the process as 5 separate steps. First new elements may be introduced into the set of contribution c^i strategies. Next, the beliefs are updated via Eqn. (4.4) and a new average belief is produced $\bar{b}_{t+1}^i$. The expected payoff based on $\bar{b}_{t+1}^i$ for each of the J contribution choices is created. Then replication takes place to create a new remembered strategy set A_{t+1}^i. Finally, the probability measure ψ_{t+1}^i is calculated and used to select the strategy to be played next period.

The first step is referred to as the *Experimentation* stage. New strategies are introduced with a probability of ρ. If strategy j is selected for experimentation, then the new contribution strategy will be $\hat{c}_j^i = round(c_j^i + N(0, \theta))$ where $N(0, \theta)$ is the normal distribution with a mean of zero and a standard deviation of θ. If $\hat{c}_j^i$ is less than zero or greater than w, then $\hat{c}_j^i$ is redrawn using $N(0, \theta)$ until $0 \leq \hat{c}_j^i \leq w$.

Belief Updating is the next step of the algorithm. The beliefs are updated using the adaptive expectations discussed previously with $b_{j,t+1}^i = (1-\lambda) * b_{j,t}^i + \lambda * \mu_t^i \; \forall j$. The final portion is to create a new average belief $\bar{b}_{t+1}^i = (1/J) * (\sum_{k=1}^J b_{k,t+1}^i)$ for use in calculating the expected payoffs. Thus the average belief includes both the average contribution of others (i.e. the static expectations) as well as the past beliefs (i.e. an anchoring bias).

Calculating the *Expected Payoff* is the third step of the algorithm. Each of the J possible contribution choices is calculated as $E_{t+1}[u_j^i(c_j^i|\bar{b}_{t+1}^i)] = c_j^i * [(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{i*}(\frac{N-1}{N})] + \bar{b}_{t+1}^i * (N-1) * [M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^{i*}}{N}] + w * (1 + \beta^i)$. The expected payoff is then used to determine which of the strategies continues to remain in the remembered strategy set as well as the probability of a strategy being selected for play in the next period.

The fourth step is called *Replication* and is used to create the set of remembered strategies in the next period, A_{t+1}^i. Each strategy has an equal chance of being selected from the

[15] If there was no elicitation of beliefs, then no rounding occurred

[16] See Ross et. al. (1977) for an example from the psychological literature and Englemann & Strobel (2000) and Blanco et. al. (2011) for examples from economics

current set of remembered strategies for use in a pairwise competition based on the expected payoff. For strategies k and $\ell \in J$ selected from A_t^i, strategy k will be copied into A_{t+1}^i whenever $E_{t+1}[u_k^i(c_k^i|\bar{b}_{t+1}^i)] > E_{t+1}[u_\ell^i(c_\ell^i|\bar{b}_{t+1}^i)]$ while strategy ℓ will be copied into A_{t+1}^i whenever $E_{t+1}[u_\ell^i(c_\ell^i|\bar{b}_{t+1}^i)] > E_{t+1}[u_k^i(c_k^i|\bar{b}_{t+1}^i)]$. If $E_{t+1}[u_k^i(c_k^i|\bar{b}_{t+1}^i)] = E_{t+1}[u_\ell^i(c_\ell^i|\bar{b}_{t+1}^i)]$, then each strategy is equally likely to be copied into A_{t+1}^i. This step is repeated J times to create a new remembered strategy set.

The final step is called *Selection* and determines which of the rules in the remembered strategy set A_{t+1}^i will be played in the next period. Selection is based on the same probability ψ^i used in initializing the algorithm. This means the probability strategy j will be used is equal to $\psi_{j,t+1}^i = \frac{E_{t+1}[u_j^i(c_j^i|\bar{b}_{t+1}^i)]}{\sum_{k=1}^J E_{t+1}[u_k^i(c_k^i|\bar{b}^i)]_{t+1}}$. Once the strategy is selected, the strategy will be played in period $t+1$ and the evolutionary process will begin again.

4.2.3 Simulation Results

The results from simulating the algorithm are the focus of this section. The algorithm itself utilizes the same set of parameters used in Arifovic & Ledyard (2012) with the IEL $(J, \rho, \theta) = (100, 0.033, 0.1 * w)$ and the ORP $(P, B, G) = (0.48, 22, 8)$. Each of the simulations are run using MatLab R2010a ©for 500 iterations using the same number of periods and observations implemented in the original experiment. To provide a consistent measure, the normalized mean squared error will be utilized with the mean squared error or MSE $= (\bar{c}_{Exp}^{All} - \bar{c}_{IEL}^{All})^2 + (\bar{c}_{Exp}^{L3} - \bar{c}_{IEL}^{L3})^2$ where $\bar{c}_{Exp}^{All}$ is the average contribution observed in the experiment over all 10 periods, IEL indicates the results from the simulations, and $L3$ indicates the average outcome of the last 3 periods using 10 tokens as the baseline[17]. This will be normalized based on the number of different experiments conducted with $NMSE = \sqrt{MSE/(2 * R)}$ where R is the number of experiments. Finally, the algorithm is implemented using the endowment w specified in each experiment[18].

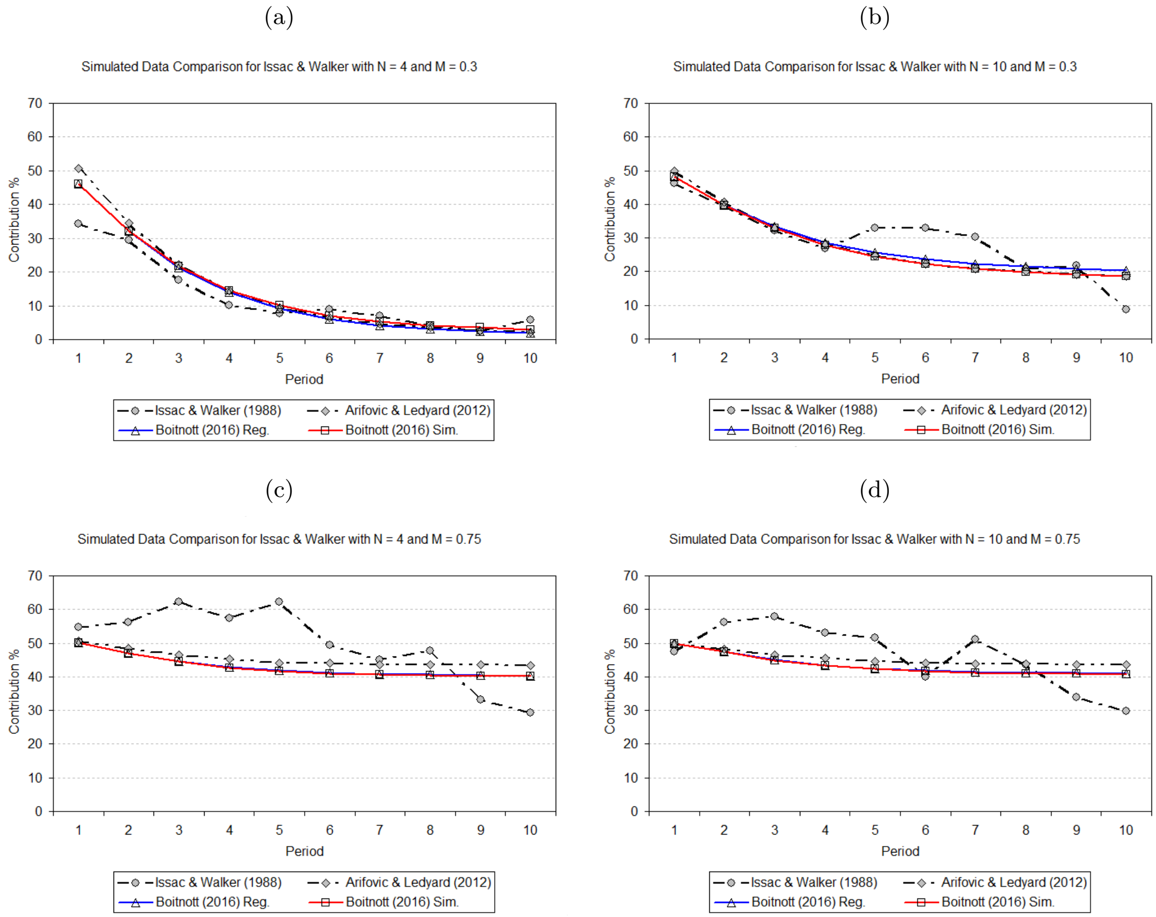
The algorithm is first applied to the Issac & Walker data. As previously discussed, the experiment consisted of a 2 by 2 design with $N = [4, 10]$ and $M = [0.3, 0.75]$. The algorithm is run using the λ found by both the regression and the simulation analysis. The $NMSE$ from the regression λ is 0.387 and the $NMSE$ from the simulation λ is 0.368, while Arifovic & Ledyard achieved a $NMSE$ of 0.341[19]. Figures 4.1a, 4.1b, 4.1c, and 4.1d show the average outcome of the Issac & Walker experiments, the Arifovic & Ledyard simulation, and the outcome of the belief-based IELORP* algorithm based on the two different values of λ. The Sim outcomes are produced by the λ determined by the simulations while the

[17]the contribution is normalized so that $\bar{c} \in [0, 10]\, \forall c$

[18]The ORP parameters found in Arifovic & Ledyard (2012) was based off Issac & Walker (1988) using $w = 10$ for all four experimental settings, yet the endowment w differed based on the experimental settings as shown in Table 1 of Issac & Walker (1988) on page 188.

[19]The original code of the algorithm producing the Issac & Walker data was released, but produces a $NMSE = 0.424$ and using either λ gives a lower $NMSE$. The difference between what is reported in the paper and the higher value here is likely the result of using a newer version of MatLab.

Figure 4.1: IEL Models Compared to Isaac & Walker(1988) Data

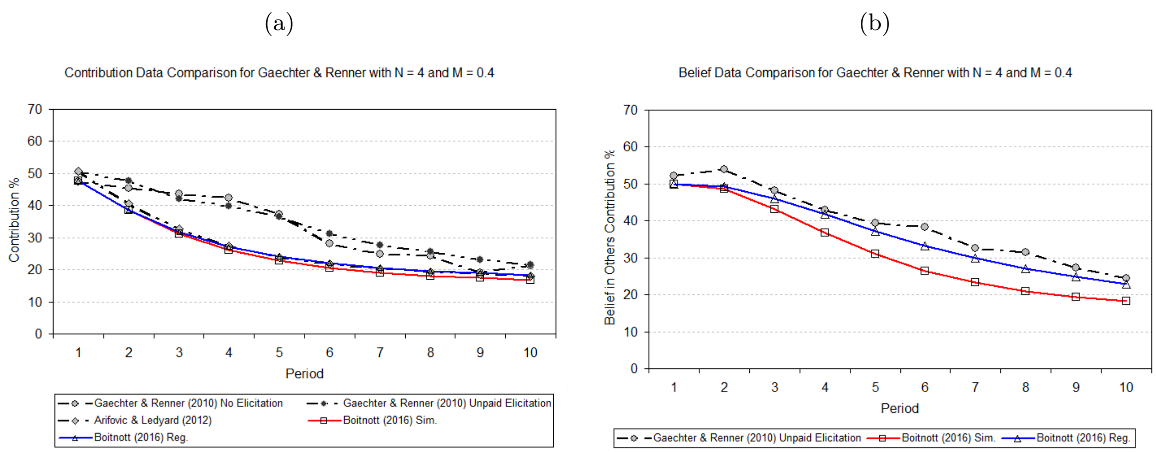


Reg outcomes are produced by the λ determined by the regression. Figures 4.1a and 4.1b show that the introduction of beliefs, a different initialization of the algorithm, and limiting the contribution to the indivisible units used in the experimental setting has a minimal influence for $M = 0.3$. When $M = 0.75$, the belief-based IELORP* results in a steeper decline in contributions compared to the original as shown in Figures 4.1c and 4.1d.

To reduce the number of differences between the belief based IELORP* and the original code, the simulations are also run using $w = 10$ as done in the original code for all of the experiments of Issac & Walker. This change has very little influence on the $NMSE$ with the value falling to 0.381 using the λ from the regression and no change in $NMSE$ using the λ from the simulations. In both cases, there is little difference between the $NMSE$ of Arifovic & Ledyard and those when the algorithm utilizes beliefs via adaptive expectations. Examining how the algorithm performs relative to the Issac & Walker data is a test to ensure that updating beliefs via adaptive expectations does not undermine the outcomes produced by the algorithm. While the accuracy fell by approximately 13% or 7% (depending on the choice of λ), the overall $NMSE$ remains relatively small suggesting a close approximation of the aggregate result.

The next experiment examined is those conducted by Gächter & Renner (2010). As previously mentioned, the experiment consisted of three different treatments with two of the treatments relevant to this section. The control treatment had four players participate in a LPGG with M = 0.4. The second treatment relevant here had the player participate in the same experimental setting, but also elicited their beliefs about the contribution of others but did not include a monetary incentive via a proper scoring rule. The contribution choices made in these two experimental settings were statistically indistinguishable allowing for the outcomes of the algorithm to be compared with two sets of observations (35 groups in total[20]). The original IELORP* code is implemented along with the belief-based IELORP* to examine how the various simulations perform relative to the data. Because the original IELORP* does not include a system of beliefs, the $NMSE$ cannot be calculated. The values from the two treatments are compared separately to determine the $NMSE$.

Figure 4.2: IEL Models Compared to Gächter & Renner (2010) Data



Unlike the Issac & Walker outcomes, the choice of λ appears to have a large influence on the $NMSE$ for this data. The original IELORP* produces an NMSE of 0.5567 while using the regression value of λ produces an almost identical $NSME$ of 0.5696. Additionally, comparing the outcome against beliefs results in a $NMSE$ of 0.2803. The simulation λ performs much worse than the outcome from the λ obtained by regression resulting in a $NMSE$ of 0.6818 relative to contribution and 0.7714 relative to stated beliefs regarding how much others will contribute. Figures 4.2a and 4.2b show the outcomes of the simulations relative to the experimental data for the contribution choices and unpaid beliefs about how much others will contribute, respectively. While the choices of contribution made using

[20] Issac & Walker had 24 observations over 4 treatments.

the simulations are relatively similar regardless of the choice of λ, the stated beliefs differ significantly depending on how the adaptive expectations are implemented.

The λ obtained via the regression analysis of Eqn. (4.4) produces outcomes that more closely matches the Gächter & Renner data rather than the λ produced via the simulation analysis. The primary difference between the two was the inclusion of individual fixed effects to remove the variation attributable to individual idiosyncratic behavior. Thus it appears that the regression analysis produces a more reliable picture of behavior than the simulation analysis. Even though the Issac & Walker data are matched by the λ from the simulation more closely, the lack of information on beliefs does not allow for a secondary check. Additionally, the lack of LPGG experiments with participants providing un-incentivized beliefs makes further testing of which model provides a better fit an open question requiring more data before reaching a conclusion.

4.3 Incentivized Beliefs

When beliefs are elicited using a monetary reward, agents are now required to make two choices to determine their payoff in each stage. The payoff an agent receives in each stage still depends on the per capita rate of return M and the choice of contribution by each agent c^i as well as the elicited beliefs b^i. While the previous section had agents with a set of beliefs that had little relevance to their overall payoff, in this section the beliefs are instrumental in determining the potential payoff an agent receives. At this point we can add to the list of stylized facts by noting that the *overly optimistic* beliefs are seen in all of the games when beliefs are incentivized. In other words, the experimental subjects provided higher stated beliefs than contribution choices on average. Given the added benefit of reflecting on beliefs, this chapter examines three possible methods players may take to evaluate beliefs. The first extends the adaptive beliefs previously discussed by adjusting λ to the Gächter & Renner (2010) data when beliefs are monetized. The second method continues to assume the agents have adaptive expectation, but adds a weight to the payoff from stated beliefs to account for the changes in the LPGG payoffs from using other-regarding preferences. The third method focuses on the fact that elicited beliefs could be a strategic concern suggesting beliefs should be a separate strategy from contribution choice in the LPGG. Given the Gächter & Renner data supports the idea that the agents increase their contributions when beliefs are incentivized while the Croson (2000) data showed a fall in contribution with incentivized beliefs supports the idea that beliefs being treated strategically might fit the data better than the adaptive expectations. Since each experiment relies on a different (though generally similar) method to incentivize beliefs, each of them will be briefly reviewed with a more detailed list of the exact payoff functions provided in Appendix C.

Each of the experiments utilizes a different (though similar "proper" scoring rule) arrangement to determine the payoff an agent receives from predicting the behavior of other agents in the experiment. There are three major observable differences: relative size of payoff versus endowment, predictions based on average versus aggregate contribution choice, and continuous versus discontinuous payoffs for predictions closer to actual outcomes. Additionally, different experiments elicit beliefs at different points and may result in different behavioral outcomes. The size of the payoff from the correct predictions relative to the size of the endowment may also play a significant influence on behavior and lead to hedging behavior when the payoff for correct predictions are large[21]. Furthermore, agents attempting to predict the aggregate behavior of others may find this to be a more complicated process if attempting to predict the behavior of each individual as opposed to the average behavior.

Table 4.2: Scoring Rule Comparisons

Experiment	F & G	G & R	C	N et al	S	C & P
Prediction Type	Average	Average	Aggregate	Aggregate	Average	Average
Number of Choices	21	2101	76	101	101	11
Order of Prediction	Simul.	Simul.	Prior	Simul.	Simul.	Prior
Endowment	20	20	25	50	10	10
Max Payoff	3	20	25	25	2	20
Ratio Best/2nd Best	1/3	1/2	1/2	1/25	1/10	1/100

Table 4.2 provides an overview of the differences in the scoring rules used in the various experimental settings.[22] The information includes the experiments of Fischbacher & Gächter (2010, hereafter F & G), Gächter & Renner (2010, hereafter G & R), Croson (2000, C in Table 4.2), Neugebauer et al (2009, hereafter N et al), Smith (2013, S in Table 4.2), and Chaudhuri & Paichayontvijit (2010, hereafter C & P). The experiments are broken down into the type of prediction that is made (i.e. an Average or Aggregate guess of how much the other participants contributed in the experiment), the total number of choices available to the agents, the order in which the prediction is made relative to the contribution, the total endowment of tokens (or laboratory currency) available to contribute, the maximum payoff achievable from accurately predicting the contribution of others, and finally the ratio between achieving the first and second best payoffs any participant could achieve. The experimental setting of F & G and Smith are designed to avoid the hedging problem. Furthermore C & P did not elicit beliefs every period in any of the treatments. While these choices make sense in the context of their experimental investigation, I am interested in the behavior of agents when the possibility of hedging exists and when belief elicitation takes place every period.

[21]There may be more reasons than just a large payoff from prediction relative to endowment for hedging, see Blanco et al (2010) or Armantier & Treich (2013) for discussions of hedging with belief elicitation.

[22]The mathematical formulas used for each scoring rule are given in Appendix C.

4.3.1 Adjusting the Algorithm

The payoff function being used for the algorithms needs to be adjusted to take into account the payoff agents receive from accurately predicting the contribution made by the other players. Let Ψ be the payoff an agent receives from the accuracy of predicting the contribution of others based on the payoff functions shown in Appendix C. This will allow for a player to get an expected payoff as a function of $\Psi(b^i|\bar{b}^i)$ using $E_{t+1}[u^i_j(c^i_j, b^i_j|\bar{b}^i_{t+1}, \Psi)]$. There are two alternative adjustments to the algorithm made in this section based on the adaptive expectations previously implemented with slight modifications to the algorithm. First, I use the G & R data to estimate a new value of λ to determine if this is sufficient to adjust the algorithm to match data beyond G & R. Next, the algorithm is adjusted by altering the weight of payoff $\Psi(b^i|\bar{b}^i)$ for agents with other-regarding preferences since the payoff from the contribution choice portion of the expected payoff is generally an order of magnitude larger.

This version of the algorithm will be referred to as OLS in the next section. First, the value of λ is re-estimated via OLS using $(\mathrm{E}[b^i_{t+1}] - \mathrm{E}[b^i_t]) = \lambda * (\mu^i_t - \mathrm{E}[b^i_t]) + \delta^i + e$, where δ^i is an individual fixed effect for the G & R data when belief elicitation is monetized. This results in $\lambda = 0.392$ with a standard error of 0.048. The second adjustment is to utilize $E_{t+1}[u^i_j(c^i_j, b^i_j|\bar{b}^i_{t+1}, \Psi)] = \Psi(b^i_j|\bar{b}^i) + c^i_j * [(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{i*}(\frac{N-1}{N})] + \bar{b}^i_{t+1} * (N-1) * [M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^{i*}}{N}] + w * (1+\beta^i)$ with $\gamma^{i*} = \gamma^i$ if $\overline{\pi} \geq \pi^i + \Psi^i$ or 0 otherwise. Unsurprisingly this results in a very close fit to the G & R data despite the lack of any additional changes to the algorithm, but fails to produce results consistent with the Croson or N et al data. Further discussion and analysis of these results can be found in the next subsection.

The alternative adjustment is to maintain the same value of λ found in Section 2, but add a weight to $\Psi(b^i|\mu^i)$ to take into account the large values agents with other-regarding preferences can have for $c^i_j * [(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{i*}(\frac{N-1}{N})] + \bar{b}^i_{t+1} * (N-1) * [M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^{i*}}{N}] + w * (1+\beta^i)$ Let $\mathcal{I}^i$ be an indicator variable taking a value of 1 if the agent has other-regarding preferences and 0 otherwise. Then let $\hat{\Psi}(b^i|\hat{b}^i) = (1-\mathcal{I}^i) * \Psi^i(b^i|\hat{b}^i) + \mathcal{I}^i * (\Psi^i(b^i|\hat{b}^i))^{\phi}$. Thus the expected payoffs for agents will be: $E_{t+1}[u^i_j(c^i_j, b^i_j|\bar{b}^i_{t+1}, \hat{\Psi})] = \hat{\Psi}(b^i_j|\bar{b}^i) + c^i_j * [(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{i*}(\frac{N-1}{N})] + \bar{b}^i_{t+1} * (N-1) * [M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^{i*}}{N}] + w * (1+\beta^i)$ with $\gamma^{i*} = \gamma^i$ if $\overline{\pi} \geq \pi^i + \Psi^i$ or 0 otherwise. The value of ϕ is then estimated via a grid search trying to find the lowest NMSE of contribution choice using data from both G & R and Croson with the remainder of the algorithm being unadjusted relative to the previous section. The grid search occurs for ϕ between 1 and 101 increasing in increments of 10 with $\phi = 81$ producing the lowest NMSE across the two sets of data. The overall simulated outcomes are improved for Croson and N et al, but results in a worse fit for the G &R data. This version of the algorithm will be referred to as $\phi = 81$ when discussing the results of the algorithm.

One drawback to this approach is that the beliefs and the contribution choice are tied together as a single strategy. While the beliefs are no longer entirely "un-reflected" since they are now directly evaluated as part of the expected payoff, the beliefs are updated autonomously with the expected payoff still relying on $\bar{b}^i$. While this allows for a relatively straightforward application, the model is unable to properly account for all experimental design choices (i.e. the order of stated belief vs. contribution choice elicitation)[23].

4.3.2 An Alternative Approach: "Independent" Beliefs

This algorithm does not rely on the adaptive expectations, but instead utilizes $\mu^i = \sum_{j \neq i} c^j$ to determine both the choice of contribution and the stated belief of an agent during the replication portion of the algorithm. This allows for the strategy set of contributions and elicited beliefs[24] to be separated to take into account the order in which choices are made. This allows the agent to be strategic about their beliefs in ways the two models of adaptive expectations do not. In other words, the remembered strategy set will be $A_t^i = [[c_t^i], [b_t^i]]$ where the contribution choice and stated beliefs are independent strategies instead of $A_t^i = [(c_t^i, b_t^i)]$ where the contribution choice and stated beliefs are part of the same strategy. This separation allows for the stated belief (contribution choice) to be selected using the contribution choice (stated belief) determined to be best. The elicited beliefs continue to be based on the expectation of an agent about the average (or aggregate) contribution choice of the other agents in the game. The general steps of the algorithm will remain unadjusted from the description in Section 2.

This model is closer to the original IELORP*, but does rely on a different initialization and takes the discrete choices allowed by the experiment seriously. The first adjustment is to allow both the choice of contribution and elicited beliefs undergo *experimentation* independently. Furthermore, instead of *expected payoffs* for evaluation, there are two different *forgone utility functions* to evaluate the sets of strategies individually using $u(c^i|\mu^i)$ from Eqn. 4.3 to evaluate the contribution choices and $\Psi(b^i|\mu^i)$ to evaluate the stated beliefs for creating the new remembered strategy set A_{t+1}^i via independent *replication* for contribution choice c^i and elicited beliefs b^i.

Selection requires the most adjustment to the algorithm requiring a two step process. This process must take into account the order the agents are presented with information as shown in Table 4.2. The first step is to select a contribution choice or elicited belief to be used during that period using the normal biased roulette wheel selection method. Once the algorithm makes this choice, the selected c_j^i (or b_j^i) is used to create a new set of forgone utility functions using Eqn. 4.3. These new forgone utility functions are used to create a

[23]If the beliefs and contribution choices are treated as separate objects (similar to the method taken in the alternative approach), then the simulated beliefs begin to drift further from the experimental data with some replications producing beliefs that slope upward and contribution choices that are nearly a horizontal line.

[24]I will use elicited beliefs and stated beliefs interchangeably in this section.

new set of probabilities to implement the biased roulette wheel selection method. Thus if the stated beliefs are elicited prior to the contribution choice, then belief $b^i_{j,t+1}$ is selected with probability $\psi_\Psi(b^i_{j,t+1}) = \Psi(b^i_{j,t+1}|\mu^i_t)/(\sum_{k=1}^J \Psi(b^i_{k,t+1})$. The selected belief $\hat{b}^i$ is then used to determine the contribution choice $c^i_{j,t+1}$ which is selected with probability $\psi_u(c^i_{j,t+1}) = u(c^i_{j,t+1}|\hat{b}^i)/(\sum_{k=1}^J u(c^i_{k,t+1}|\hat{b}^i))$. Alternatively, if the beliefs are elicited concurrently with the contribution choice, then contribution choice $c^i_{j,t+1}$ is selected with probability $\psi_u = u(c^i_{j,t+1}|\mu^i_t)/(\sum_{k=1}^J u(c^k_{j,t+1}|\mu^i_t))$. The stated belief $b^i_{j,t+1}$ is selected using contribution choice $\hat{c}^i$ with probability $\psi_u = u(b^i_{j,t+1}|\hat{c}^i)/(\sum_{k=1}^J u(b^k_{j,t+1}|\hat{c}^i))$.

The final change adjusts how the algorithm is *initialized* to take into account the above selection method. Since there is no μ^i, the algorithm selects a belief or contribution at random depending on the order of elicitation. Then the stated belief $\hat{b}^i$ or contribution choice $\hat{c}^i$ is used in creating the probability ψ_u to select the other.

4.3.3 Results from the Algorithms

This section is broken down into two parts. It begins by giving a simple analysis of how effectively each algorithm performs using the same NMSE implemented in Section 2. In nearly every case, the algorithms are less effective matching the data than in the case where beliefs are not incentivized. After examining the NMSE, I explore two other important elements mentioned in the introduction. All of the experimental data points to beliefs on average being optimistic relative to the contribution level. This behavior is displayed by all of the algorithms implemented. This section concludes by showing that none of the algorithms is able to match the contradictory experimental data of Croson and G & R.

Examining behavior via the NMSE

Table 4.3 summarizes the NMSE from using each of the algorithm adjustments with each algorithm being differentiated by referring to the first as OLS to indicate the change to λ from the OLS estimate, $\phi = 81$ to indicate the weight added to the payoff of beliefs Ψ, and the third as $A^i = [[c^i], [b^i]]$ to indicate the separation of the contribution choice and the stated belief. Additionally, the NMSE was calculated for both the contribution choice and the stated beliefs to gain a better understanding of how each is influenced by the adjustments individually. As a reminder, the NMSE is calculated using the mean over all periods and the last three periods of either the contribution or the belief. Furthermore, the NMSE is all calculated as if the endowment was 10 as was done in Arifovic & Ledyard (2012). This allows an easier comparison since the base remains the same across simulations.

The OLS algorithm used data from G & R to estimate the new value of λ and unsurprisingly the G & R data have the closest fit. The OLS produces the worst fit for both the Croson data and the N et al data. The choice of λ moved closer to the λ found in Section 2 produced via the simulations giving a greater weight to the behavior of other players

Table 4.3: NMSE for Incentivized Belief Elicitation

	Contribution NMSE		
	G & R	Croson	N et al
OLS	0.2603	3.0498	2.0109
$\phi = 81$	1.3036	1.8760	0.7096
$A^i = [[c^i], [b^i]]$	2.2676	1.6395	0.4299
	Belief NMSE		
	G & R	Croson	N et al
OLS	0.1886	2.4313	1.4439
$\phi = 81$	0.7609	1.6991	0.5395
$A^i = [[c^i], [b^i]]$	1.3958	1.3052	0.6007

relative to the agents own initial beliefs. The main reason for the poor performance of the simulation relative to the data is the Ψ portion of the payoff has almost no influence on what strategy is selected for the agents with other-regarding preferences since it is generally dwarfed in scope. This results in only agents without other-regarding preferences being influenced.

The $\phi = 81$ algorithm used data from both G & R and Croson to determine the best value of ϕ. The choice of ϕ allowed those agents with other-regarding preferences to be influenced by producing payoffs to ϕ that were comparable to the influence of the other-regarding preferences. The result is an improvement in the fit of the simulations relative to the data for both Croson and N et al, but the G & R data has a decrease in the fit especially in comparison to the outcome produced in OLS simulation.

The $A^i = [[c^i], [b^i]]$ algorithm improves the fit for both Croson and N et al while, the fit for G & R is significantly worse. One beneficial element is that no additional fitting or parameter value estimation is needed to construct this model. This suggests that there is some influence to the elicitation ordering of beliefs and contributions at least when the stated belief is focused on the aggregate as opposed to the average.

To get a better understanding of the influence of the different methods of eliciting beliefs, more information is required than the three studies currently available. Using adaptive expectations to determine the beliefs of the agents does not appear to be sufficient to explain behavior except in the case of G & R, but this is not surprising since the G & R data were integral in selecting the λ or ϕ variables. The assumption that the agents select either the contribution (or belief) independently and then use this to determine the belief (or contribution) also does not appear to provide a sufficient explanation.

Analysis of Simulations Relative to Empirical Elements

Beyond simply matching the data, the goal of this section was to determine if the adaptive beliefs or the strategic beliefs were sufficient to understand more behavior in the experiments. One of the important elements mentioned is the *overly optimistic beliefs* observed in all of the experimental data. Table 4.4 shows the average contribution and belief for each of the algorithms as a percentage of the endowment. An additional point of focus is to determine if any of the simulations is able to mimic the contradictory contribution behavior displayed by G & R versus Croson when beliefs are incentivized. None of the algorithms investigated is able to match this outcome.

Table 4.4: Average Behavior in Incentivized Belief Elicitation

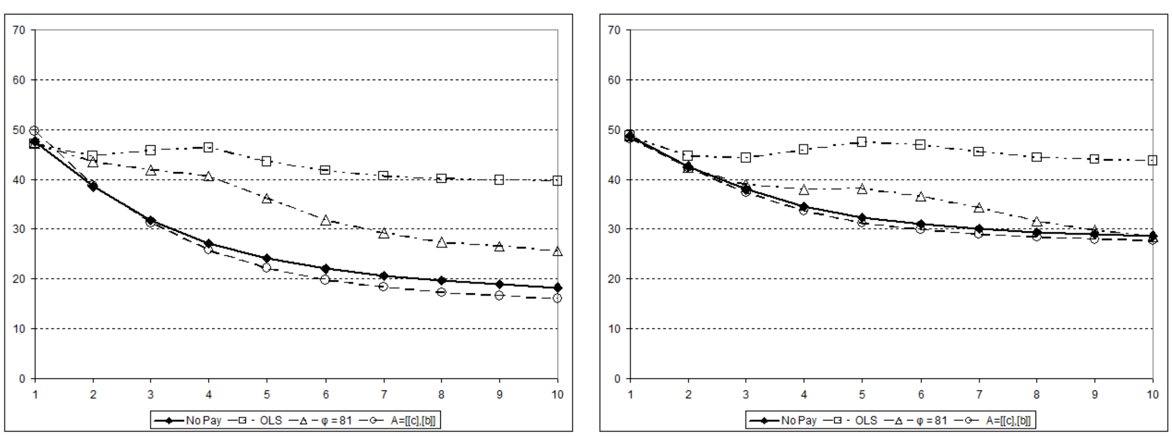
	Average Contribution		
	G & R	Croson	N et al
OLS	42.95%	45.56%	42.57%
$\phi = 81$	35.02%	36.69%	31.34%
$A^i = [[c^i], [b^i]]$	25.58%	33.54%	29.19%
	Average Belief		
	G & R	Croson	N et al
OLS	45.37%	47.00%	45.09%
$\phi = 81$	41.75%	42.47%	38.89%
$A^i = [[c^i], [b^i]]$	37.30%	38.53 %	39.75%

N et al show that beliefs are *overly optimistic* since the average of stated beliefs is generally higher than the average contribution in a period. Each of the algorithms regardless of the method used replicates this result with the average stated belief being above the average contribution in a period as shown in Table 4.4. The beliefs using $\phi = 81$ generally gives the closest match to the experimental data, while the OLS provided beliefs that were only slightly higher than the average contribution an outcome only consistent with the G & R experimental data.

Figure 4.3: Effects of Incentivized Beliefs on Behavior

(a) Simulation: G & R

(b) Simulation: Croson



Figures 4.3a and 4.3b show how contribution changes for the simulation. To be consistent with the experimental data, one of the algorithms would need to have data above the “No Pay” outcome in Figure 4.3a and below the “No Pay” outcome in Figure 4.3b. None of the algorithms comes close to being able to match this sort of behavior.

4.4 Conclusion

This chapter provides an extension of the IELORP* model by restricting the algorithm to the same discrete choices available to experimental subjects and introducing non-myopic beliefs into the model. Adaptive expectations are introduced into the model with the weight used to update beliefs estimated using Gächter & Renner (2010) data. The introduction of beliefs allows for an updating of how the algorithm is initialized. The paper then demonstrates how the algorithm continues to fit the contribution choice data from a standard LPGG while using discrete choices and adaptive expectations without needing to re-estimate the parameter values of the model. Allowing the algorithm to have more sophisticated beliefs creates an opportunity to investigate if the model is able to capture the behavior of agents when belief elicitation is incentivized. The model is only able to match the behavior of the experimental subjects from Gächter & Renner suggesting a limit of the model.

While the adaptive expectations work sufficiently to capture the average behavior in the experiments without incentivized beliefs, this form of beliefs lacks the ability to capture the level of variance observed in experiments. The adaptive expectations are also not sufficiently sophisticated/forward-looking to have simulated agents increase their contribution choice in the hopes of convincing others to contribute more as modeled by Wendel & Oppenheimer (2010). Incentivized belief elicitation has proven to be more problematic for the adaptive expectations given the limited number of data sources available; however, using myopic agents with strategies that take into account the ordering of stated beliefs and contribution choices in the experiment does not appear to be sufficient to explain behavior when beliefs are monetized via a proper scoring rule.

Future research should consider other methods for implementing both the adaptive expectations and the ability to take into account the ordering of elicitation. The initialization of the algorithm could use a greater level of sophistication as the random assignment of beliefs and contribution choices does not appear to be sufficient. In reality, a more complex set of initial behavior should take into account the M and N in ways that are consistent with the accepted stylized facts.

Furthermore, additional experiments examining the influence of proper scoring rules eliciting beliefs as averages versus aggregate expectations as well as the influence of continuous versus discontinuous payoffs. Of course, given the claim of other-regarding preferences that will influence behavior based on the actual actions of others, a pre-experiment using the strategy method implemented by Fischbacher & Gächter (2010) could further the understanding of the relationship between contribution choices and incentivized belief elicitation.

Chapter 5

Conclusion

This thesis has focused on the application of the Individual Evolutionary Learning (IEL) model. The question of external validity is a constant issue for learning models, but this thesis demonstrates a wide array of games where the IEL is able to match the experimental data with a consistent set of parameter choices. Additionally, each of the chapters attempts to mimic the experimental design as closely as possible to provide the algorithm with an approximate match of available behavior. In Chapter 2, I demonstrated how the IEL was better able to match most of the individual level behavior exhibited in the games being explored. Additionally, for the games of complete information, the IEL clearly outperformed any of the other algorithms examined. I worked with Jasmina Arifovic and John Duffy in Chapter 3 to demonstrate the behavior of the IEL model in games focused on correlated equilibria. The IEL almost perfectly matched experimental data from a Battle of the Sexes Game, but had more difficulty matching outcome from a Chicken Game. Chapter 4 shows an extension of the IEL with Other-Regarding Preferences where beliefs modeled as adaptive expectations and the available actions match the limited choices of the experimental data. Despite a wide number of adjustments to the algorithm, the simulated data still closely matched the results of the Linear Public Goods Game when either belief elicitation is not monetized or no belief elicitation occurs. The algorithm is generally unable to match the outcomes when beliefs are elicited using a monetary incentive. The flexibility and transferability of the IEL suggests a rather robust model that has a great deal of potential to be applied more widely than is currently being done.

While not the focus on any of the chapters in this thesis, an important area of study to gain a deeper understanding of how the IEL models behavior is to create ex-ante predictions of behavior for comparison. This allows for two general comparisons that may illuminate a greater understanding of the algorithm. First, it allows the experiment to be designed to match the information that would be available to the IEL. Since the algorithm creates certain restrictions on information, this would allow the experimentalist an opportunity to gain a greater understanding of information the experimental subjects have. Second, any adjustments needed for the algorithm to be able to more closely match the experimental data would allow for a comparison with the ex-ante predictions. The ex-post versus ex-ante predictions would provide a clear method of comparing the elements of the algorithm that needed adjustments.

Bibliography

[1] Andreoni, James. "Why Free Ride? Strategies and Learning in Public Goods Experiments." *Journal of Public Economics*, 37 (1988): 291-304.

[2] Andreoni, James and Rachel T.A. Croson. "Partners Versus Strangers: Random Rematching in Public Goods Experiments." *Handbook of Experimental Results*, 1 (2008): 776-783.

[3] Arifovic, Jasmina, and John Ledyard. "Scaling Up Learning Models in Public Good Games." *Journal of Public Economic Theory*, 6 (2004): 203-238.

[4] Arifovic, Jasmina, and John Ledyard. "Call Market Book Information and Efficiency." *Journal of Economic Dynamics and Control*, 31 (2007): 1971-2000.

[5] Arifovic, Jasmina, and John Ledyard. "A Behavioral Model for Mechanism Design: Individual Evolutionary Learning." *Journal of Economic Behavior and Organization*, 78 (2010): 374-395.

[6] Arifovic, Jasmina, and John Ledyard. "Individual Evolutionary Learning, Other-Regarding Preferences, and the Voluntary Contributions Mechanism." *Journal of Public Economics*, 96 (2012): 808-823.

[7] Arifovic, Jasmina, Joshua F. Boitnott, and John Duffy. "Learning Correlated Equilibrium: An Evolutionary Approach." *Working Paper* (2015).

[8] Arifovic, Jasmina, Joshua F. Boitnott, and John Duffy. "Learning Correlated Equilibrium: An Evolutionary Approach." forthcoming *Journal of Economic Behavior and Organization* (2016).

[9] Armantier, O. and N. Treich. "Eliciting Beliefs: Proper Scoring Rules, Incentives, Stakes, and Hedging." *European Economic Review*, 62 (2013): 17-40.

[10] Arthur, W. Brian. "On Designing Economic Agents that Behave like Human Agents." *Journal of Evolutionary Economics*, 3 (1993): 1-21.

[11] Aumann, Robert J. "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics*, 1 (1974): 67-96.

[12] Aumann, Robert J. "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica*, 55 (1987): 1-18.

[13] Axelrod, Robert. *The Evolution of Cooperation*, Basic Books, Inc., 1984.

[14] Bayer, Ralph-C., Elke Renner, and Rupert Sausgruber. "Confusion and Learning in the Voluntary Contributions Game." *Experimental Economics*, 16 (2013): 478-496.

[15] Ben-Porath, Elchanan. "Correlation without Mediation: Expanding the Set of Equilibrium Outcomes by "Cheap" Pre-play Procedures." *Journal of Economic Theory*, 80 (1998): 108-122.

[16] Blanco, Mariana, Dirk Engelmann, Alexander Koch, and Hans-Thep Normann. "Belief Elicitation in Experiments: Is There a Hedging Problem?" *Experimental Economics*, 13 (2010): 412-438.

[17] Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann. "A Within-Subject Analysis of Other-Regarding Preferences." *Games and Economic Behavior*, 72 (2011): 321-338.

[18] Boitnott, Joshua F. "The Behavior of Learning Algorithms in Experimental Games with Congestion." *Working Paper*, (2015).

[19] Bone, John, Michalis Drouvelis and Indrajit Ray. "Coordination in 2×2 Games by Following Recommendations from Correlated Equilibria." *University of Birmingham Economics Discussion Paper* (2013): 13-04.

[20] Branch, William A. "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *The Economic Journal*, 114 (2004) 592-621.

[21] Burton-Chellew, Maxwell N., Claire Mouden, and Stuart A. West. "Conditional Cooperation and Confusion in Public-Goods Experiments." *Proceedings of the National Academy of Sciences*, 113 (2016): 1291-1296.

[22] Cason, Timothy N., and Tridib Sharma. "Recommended Play and Correlated Equilibria." *Economic Theory*, 33 (2007): 11-27.

[23] Camerer, Colin, and Teck Hua Ho. "Experience Weighted Attraction Learning in Normal Form Games." *Econometrica*, 67 (1999): 827-874.

[24] Chalet, Damien and Yi-Chen Zhang. "Emergenence of Cooperation and Organization in an Evolutionary Game." *Physica A: Statistical Mechanics and its Applications*, 246 (1997): 407-418.

[25] Chaudhuri, Ananish. "Sustaining Cooperation in Laboratory Public Goods Experiments: a Selective Survey of the Literature." *Experimental Economics* 14 (2011): 47-83.

[26] Chaudhuri, Ananish, and Timud Paichayontvijit. "Does Strategic Play Explain the Decay in Contributions in a Public Goods Game? Experimental Evidence." *Working Paper*, University of Auckland, (2010).

[27] Chen, Shu-Heng. "Varieties of Agents in Agent-Based Computational Economics: A Historical and an Interdisciplinary Perspective." *Journal of Economic Dynamics & Control*, 36 (2012): 1–25.

[28] Chen, Yan and Yuri Khoroshilov. "Learning under Limited Information." *Games and Economic Behavior*, 44, No. 1 (2003): 1-25.

[29] Chmura, Thorsten and Thomas Pitz. “Minority Game: Experiments and Simulations of Traffic Scenarios.” *Bonn Econ Discussion Papers*, No. 23 (2004).

[30] Chmura, Thorsten and Thomas Pitz. “Successful strategies in repeated minority games.” *Physica A: Statistical Mechanics and its Applications*, 363, No. 2 (2006): 477-480.

[31] Chmura, Thorsten and Güth. “The Minority of Three-Game: An Experimental and Theoretical Analysis.” *Games*, 2 (2011): 333-354.

[32] Chmura, Thorsten, Sebastian J. Goerg, and Reinhard Selten. “Learning in Experimental 2x2 Games.” *Games and Economic Behavior*, 76 (2012): 44-73.

[33] Clemens, Christiane, and Thomas Riechmann. “Evolutionary Dynamics in Public Good Games.” *Computational Economics*, 28 (2006): 399-420.

[34] Cotla, Chenna R., and Ragan Petrie. “Stability of Social Preferences and Learning in Public Goods Games.” *Working Paper* (2015).

[35] Croson, Rachel T.A. “Partners and Strangers Revisited.” *Economic Letters*, 53 (1996): 25-32.

[36] Croson, Rachel T.A. “Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play.” *Journal of Economic Behavior & Organization*, 41 (2000): 299-314.

[37] Donoso, Maäl, Anne G. E. Collins, and Etienne Koechlin. “Foundations of Human Reasoning in the Prefrontal Cortex.” *Science*, 344 (2014): 1481-1486.

[38] Duffy, John, and Ed Hopkins. “Learning, Information, and Sorting in Market Entry Games: Theory and Evidence.” *Games and Economic Behavior*, 51 (2005): 31-62.

[39] Duffy, John. “Agent-Based Models and Human Subject Experiments.” *Handbook of Computational Economics*, 2 (2006): 949-1011.

[40] Duffy, John and Nick Feltovich. “Correlated Equilibria, Good and Bad: An Experimental Study.” *International Economic Review*, 51 (2010): 701-721.

[41] Duffy, John, Ernest K. Lai and Wooyoung Lim. “Coordination via Correlation: An Experimental Study.” *Economic Theory*, forthcoming (2016).

[42] Dufwenberg, Martin, and Georg Kirchsteiger. “A Theory of Sequential Reciprocity.” *Games and Economic Behavior*, 47 (2004): 268-298.

[43] Englemann, Dirk, and Martin Strobel. “The False Consensus Effect Disappears if Representative Information and Monetary Incentives are Given.” *Experimental Economics*, 3 (2000): 241-260.

[44] Erev, Ido and Alvin E. Roth. “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term.” *Games and Economic Behavior*, 8 (1995): 164-212.

[45] Erev, Ido and Alvin E. Roth. “Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria.” *The American Economic Review* 88 (1998): 848-881.

[46] Erev, Ido, Eyal Ert, and Alvin E. Roth. "A Choice Prediction Competition for Market Entry Games: An Introduction." *Games*, 1 (2010a): 117-136.

[47] Erev, Ido, Eyal Ert, and Alvin E. Roth. "A Choice Prediction Competition for Market Entry Games: Correction." *Games*, 1 (2010b): 221-225.

[48] Fehr, Ernst, and Simon Gächter. "Fairness and Retaliation: The Economics of Reciprocity." *The Journal of Economic Perspectives*, 14 (2000): 159-181.

[49] Feltovich, Nick. "Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games." *Econometrica*, 68 (2000): 605-641.

[50] Fischbacher, Urs, and Simon Gächter. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *The American Economic Review*, 100 (2010): 541-556.

[51] Forges, Françoise. "Equilibria with Communication in a Job Market Example." *Quarterly Journal of Economics*, 105 (1990): 375-398.

[52] Foster, Dean P. and Rakesh V. Vohra. "Calibrated Learning and Correlated Equilibrium." *Games and Economic Behavior*, 21 (1997): 40-55.

[53] Fudenberg, Drew and David K. Levine. "Universal Conditional Consistency." *Games and Economic Behavior*, 29 (1999): 104-130.

[54] Fuster, Andreas, David Liabson, and Brock Mendel. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives*, 24 (2010) 67-84.

[55] Gäcther, Simon, and Elke Renner. "The Effects of (Incentivizes) Belief Elicitation in Public Goods Experiments" *Experimental Economics*, 13 (2010): 364-377.

[56] Gintis, Herbert. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press, 2009.

[57] Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. "When and Why Incentives (Don't) Work to Modify Behavior." *The Journal of Economic Perspectives*, 25 (2011): 191-210.

[58] Hart, Sergiu and Andreu Mas-Colell. "A Simple Adaptive Procedure Leading to Correlated Equilibrium", *Econometrica*, 68 (2000): 1127 -1150.

[59] Ho, Teck H., Colin Camerer, and Juin-Kuan Chong. "Self-tuning experience-weighted attraction learning in games." *Journal of Economic Theory*, 133 (2007): 177–198.

[60] Holt, Charles, and Susan K. Laury. "Theoretical Explanations of Treatment Effects in Voluntary Contributions Experiments." *Handbook of Experimental Economics Results*, 1 (2008): 846-855.

[61] Isaac, R. Mark, and James M. Walker. "Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism." *Quarterly Journal of Economics*, 103 (1988): 179-199.

[62] Kahneman, Daniel and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47 (1979): 263-292.

[63] Kahnemen, Daniel. *Thinking Fast and Slow*, MacMillan, 2011.

[64] Keser, Claudia. "Voluntary Contributions to a Public Good When Partial Contribution is a Dominant Strategy." *Economic Letters*, 50 (1996): 359-366.

[65] Klumpp, Tilman. "Finitely Repeated Voluntary Provision of a Public Good." *Journal of Public Economic Theory*, 14 (2012): 547-572.

[66] Kocher, Martin G., Peter Martinsson, Dominik Matzat, and Conny Wollbrant. "The Role of Beliefs, Trust, and Risk in Contributions to a Public Good." *Working Paper*, (2011).

[67] Kuzmics, Christoph, Thomas Palfrey, and Brian W. Rogers. "Symmetric Play in Repeated Allocation Games." *Journal of Economic Theory*, 154 (2014): 25-67.

[68] Ledyard, John. "Public Goods: A Survey of Experimental Research." *The Handbook of Experimental Economics*, (1995): 111-194.

[69] Lehrer, Ehud. "A Wide Range No-Regret Theorem." *Games and Economic Behavior*, 42 (2003): 101-115.

[70] Moreno, Diego and John Wooders. "An Experimental Study of Communication and Coordination in Noncooperative Games." *Games and Economic Behavior*, 24 (1998): 47-76.

[71] Muth, John F. "Properties of Some Short-Run Business Forecasts." *Eastern Economic Journal*, 3 (1985): 200-210.

[72] Myerson, Rodger B. *Game Theory: Analysis of Conflict.* Cambridge: Harvard University Press, 1991.

[73] Neugebauer, Tibor, Javier Perote, Ulrich Schmidt, and Malte Loos. "Selfish-Biased Conditional Cooperation: On the Decline of Contributions in Repeated Public Goods Experiments." *Journal of Economic Psychology*, 30 (2009): 52-60.

[74] Nyarko, Yaw, and Andrew Schotter. "An Experimental Study of Belief Learning Using Elicited Beliefs." *Econometrica*, 70 (2002): 971-1005.

[75] Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 80 (1993): 1281-1302.

[76] Rapoport, Amnon, Tamar Kugler, Subhasish Dugar, and Eyran J. Gisches. "Choice of Routes in Congested Traffic Networks: Experimental Tests of the Braess Paradox." *Games and Economic Behavior*, 65, No. 2 (2009): 538-571.

[77] Rosenthal, Robert W."A Class of Games Possessing Pure-Strategy Nash Equilibria." *International Journal of Game Theory*, 2 (1973): 65-67.

[78] Ross, Lee, David Greene, and Pamela House. "The 'False Consensus Effect': An Egocentric Bias in Social Perception and Attribution Processes." *Journal of Experimental Social Psychology*, 13 (1977): 279-301.

[79] Ruström, E. Elisabet, and Nathaniel T. Wilcox. "Stated Beliefs Versus Inferred Beliefs: A Methodological Inquiry and Experimental Test." *Games and Economic Behavior*, 67 (2009): 616-632.

[80] Schlag, Karl H., James Tremewan, and Joel J. van der Weele. "A Penny for Your Thoughts: A Survey of Methods for Eliciting Beliefs." *Experimental Economics*, 18 (2015): 457-490.

[81] Selten, Reinhard, Klaus Abbink, and Ricarda Cox. "Learning Direction Theory and the Winner's Curse." *Experimental Economics*, 8, No. 1 (2005): 5-20.

[82] Selten, Reinhard, Thorsten Chmura, Thomas Pitz, Sebastian Kube, and Michael Schreckenberg. "Commuters Route Choice Behaviour." *Games and Economic Behavior*, 38, No. 2 (2007): 394-406

[83] Selten, Reinhard and Thorsten Chmura. "Stationary Concepts for Experimental 2x2-Games." *American Economic Review*, 98, No. 3 (2008): 938-966.

[84] Smith, Alexander. "Estimating the Causal Effect of Beliefs on Contributions in Repeated Public Good Games." *Experimental Economics*, 16 (2013): 414-425.

[85] Smith, Alexander. "Modeling the Dynamics of Contributions and Beliefs in Repeated Public Good Games." *Economic Bulletin*, 35 (2015): 1501-1509.

[86] Wendel, Stephen, and Joe Oppenheimer. "An Agent-Based Analysis of Context-Dependent Preferences." *Journal of Economic Psychology*, 31 (2010): 269-284.

[87] Wilcox, Nathaniel T., and Nick Feltovich. "Thinking Like a Game Theorist: Comment." *University of Houston Department of Economics Working Paper*, (2000).

[88] Zelmer, Jennifer. "Linear Public Goods Experiments: A Meta-Analysis." *Experimental Economics*, 6 (2003): 299-310.

Appendix A

Additional Material: Chapter 2

A.1 Additional Algorithms and Details

This section of Appedix A provides greater detail about the IEL as well as a summary of two additional learning algorithms that analyzed in this Appendix.

A.1.1 Individual Evolutionary Learning Alternative Distribution

To allow the IEL to select values equal to the boundaries, the following distribution was used:

i if $x < L$, $x = L$;

ii if $x \in [L, U]$, $x = E(x|x \in [L, U])$; and,

iii if $x > U$, $x = U$

The distribution described by elements i, ii, and iii gives the following mean and variance:

$$E(x|i, ii, iii) = L * \Phi(\tilde{L}) + \mu * (\Phi(\tilde{U}) - \Phi(\tilde{L})) - \sigma * (\phi(\tilde{U}) - \phi(\tilde{L})) + U * (1 - \Phi(\tilde{U})); \text{and,}$$

$$\begin{aligned} V(x|i, ii, iii) = {} & L * \Phi(\tilde{L}) * [L - 2(\mu * c - \sigma * d)] + U * (1 - \Phi(\tilde{U})) * [U - 2(\mu * c - \sigma * d)] \\ & -(L * \Phi(\tilde{L}) + U * (1 - \Phi(\tilde{U})))^2 + (\mu^2 + \sigma^2) * c - (\mu * c - \sigma * d)^2 - 2 * \mu * \sigma * d \\ & -\sigma^2 * (\tilde{U} * \phi(\tilde{U}) - \tilde{L} * \phi(\tilde{L})), \end{aligned}$$

where $\phi(y)$ is the value of the normal pdf at y, $\Phi(y)$ is the value of the normal cdf at y, $\tilde{U} = \frac{U-\mu}{\sigma}$ and $\tilde{L} = \frac{L-\mu}{\sigma}$, $c = \Phi(\tilde{U}) - \Phi(\tilde{L})$, and $d = \phi(\tilde{U}) - \phi(\tilde{L})$. This allows for the expected value of the IEL Alt to be smaller than the expected value produced using the IEL distribution if $\mu < \frac{U+L}{2}$ and larger than the expected value of IEL if $\mu > \frac{U+L}{2}$. The variance of the IEL Alt will be larger than the variance produced using the truncated normal distribution of the IEL.

A.1.2 The Survival Rate of Strategies via IEL Replication Method

The survival rate of a strategy is based on the design to repopulate the remembered strategy set A^i, the rule is that 2 possible strategies are selected at random with replacement from the J possible strategies and that this is done J times. I began with the simple case where ℓ is the strategy with the highest possible forgone utility and there exists i strategies identical to ℓ. The probabilities are rather straightforward as to how frequently ℓ will be selected. If I let x be the strategy selected first and y be the strategy selected second, then $p(x = \ell) = \frac{i}{J}$ and $p(y = \ell | x \neq \ell) = p(x \neq \ell)p(y = \ell) = \frac{J-i}{J} \times \frac{i}{J-1}$. Thus the probability ℓ will not be selected any given period is simply $1 - \frac{i(J-1)+i(J-i)}{J(J-1)}$. Since there are J strategies in the set of remembered strategies, this takes place J times meaning the probability ℓ will not be in next period's remembered strategy set is $(\frac{(J-i)(J-(i+1))}{J(J-1)})^J$. Taking the limit as J tends to infinity, I can determine how frequently this occurs by utilizing L'Hopital's Rule.

$$\begin{aligned}
\lim_{J\to\infty}(\frac{(J-i)(J-(i+1)}{J(J-1)})^J &= \lim_{J\to\infty}\exp[J * ln(\frac{(J-i)(J-(i+1))}{J(J-1)})] \\
&= \exp[\lim_{J\to\infty}(\frac{ln((J-i)(J-(i+1))-ln(J)-ln(J-1)}{1/J})] \\
&= \exp(-i)
\end{aligned}$$

The next case examined is where the strategy does not produce the highest forgone utility will be examined. Since there are J possible strategies, there are four possible groups. I am interested in determining the probability strategy ℓ will remain in the remembered strategy set after repopulation. Let i represent the total number of strategies identical to ℓ. When two strategies have the same forgone utility, then each strategy is equally likely to survive into the next period. Let α represent the total number of strategies that give a lower forgone utility payoff than strategy ℓ. Let ω represent the total number of strategies that will give a higher payoff than the forgone utility payoff than strategy ℓ. Thus, $J = i+\alpha+\omega$. Letting x be the strategy selected first and y be the strategy selected second, then the probability that strategy ℓ is selected and survives is:

$$p(x \in i) \cdot (p(y \in i) + p(y \in \alpha)) + p(x \in \alpha) \cdot p(y \in i)$$

Once simplified, the one period selection and survival rate for strategy ℓ is $\frac{2Ji-i(2i+2\omega+1)}{J(J-1)}$. But the probability that ℓ will be removed from the population during the J iterations is given by $(\frac{(J-i)(J-(i+1))+i(i+2\omega)}{J(J-1)})^J$. Since J is missing from the second term, there will be no difference in the probability of survival at the limit. For smaller values of J, an increase in ω will mean ℓ is less likely to survive. The effect of an increase in i is more ambiguous with the first order condition resulting in an outcome of $\frac{(2J(i-\alpha)+J)((J-i)(J-(i+1))+i(i+2\omega))^{J-1}}{(J(J-1))^J}$. This suggest that if i increases to $\hat{i}$ with $\hat{i} < \alpha$, then the probability strategy ℓ remains in the set of remembered strategies will not improve; however, if $\hat{i} > \alpha$, then the probability strategy ℓ remains in the set of remembered strategies will increase.

A.1.3 Impulse Balance Learning

The learning model creates an impulse on some action i whenever action i would have given a better payoff than the actual action j taken; however, the size and influence of the impulse depends on a transformed payoff function of the game. The algorithm has only been applied to games with 2 actions thus far in the literature. The aspiration level based on maximin payoff in games is used to transform the payoff function so long as the payoffs are > 0 since there has been no discussion of how the transformation should be made when payoffs are negative. There is a slight discussion in relation to reinforcement learning as to how this would change the payoffs; however, I assume that the entire payoff function is adjusted.

Each action i has an impulse sum $R_i(t)$ that is the sum of all impulses from action j to action i for all previous periods. Each action i is played with probability $p_i(t)$ where

$$p_i(t) = \frac{R_i(t)}{\sum_{k=1}^{K} R_k(t)}$$

with K being the total number of possible actions available to each agent. And the impulse in period t is $r_i(t)$ where:

$$r_i(t) = \begin{cases} max[0, \hat{\pi}_i - \hat{\pi}_j] & \text{if action } j \text{ is the action used} \\ 0 & \text{otherwise} \end{cases}$$

where $\hat{\pi}_i$ is the transformed hypothetical payoff from taking action i and $\hat{\pi}_j$ is the transformed payoff from using action j. The impulse sum is then updated such that $R_i(t+1) = R_i(t) + r_i(t)\ \forall i = 1, \ldots, K$. The initial impulse sum are set to zero, i.e. $R_i(1) = 0\ \forall i = 1, \ldots, K$ and the probabilities are $p_i(t) = 1/K\ \forall i = 1, \ldots, K$ whenever $\sum_{k=1}^{K} R_k(t) = 0$.

A.1.4 Experience Weighted Attraction

The Experience Weighted Attraction model was proposed by Camerer and Ho (1999). The idea is that players have some level of attraction to use some strategy in a game and that this attraction (along with the other attractions) will determine how likely the agent is to play that strategy. The algorithm depends on hypothetical forgone utilities to evaluate all possible strategies. Under certain extreme values, the EWA model is simply a version of fictitious play or reinforcement learning[1].

Each strategy (or actions) available to an agent has an attraction $Atr_j^i(t)$ where j is the strategy, i is the player, and t is the current period. Each attraction is associated with a probability $\gamma_j^i(t) = e^{\lambda \cdot Atr_j^i(t)} / \sum_{k=1}^{J} e^{\lambda \cdot Atr_k^i(t)}$ of being played. EWA allows for strategies that are not played to be evaluated; however, the strategy that is currently being used gets the full value of the outcome while the unused strategies may receive less than the full value. The algorithm works in the following way: (i) Each of the players starts with the same experience weight $N(t)$ that evolves over time according to:

$$N(t) = \rho \cdot N(t-1) + 1$$

[1]See p. 181 of Ho et al (2007) for a graphical representation.

where ρ is a depreciation rate. This leads to a discounting of the attraction as well as a weight on the hypothetical utility that a player could achieve. (ii) The weight is:

$$L_j^i(t) = \frac{\delta + (1-\delta) \cdot \mathcal{I}_j^i(s^i(t))}{N(t)}$$

where δ is the importance placed on the hypothetical payoffs and $\mathcal{I}_j^i(s^i(t))$ is an indicator variable equal to one if the j^{th} strategy was played in period t, i.e. $s^i(t) = s_j^i(t)$ was used by the player. (iii) This allows for the calculation of the new attractions:

$$Atr_j^i(t) = \phi \cdot Atr_j^i(t-1) + L_j^i(t) \cdot \pi(s_j^i|s^{-i}(t))$$

where ϕ is a discount factor. The hypothetical utility $\pi(s_j^i|s^{-i}(t)$ is based on the payoffs described in each experiment from playing strategy $s_j^i(t)$ given all other players do not change their strategies $s^{-i}(t)$ where $-i$ indicates all other players. Finally, the strategy is selected using the probability measure $\gamma_j^i(t)$ mentioned previously and is played against a random variable $\psi \sim U[0,1]$ to determine the strategy played in the next period.

This means that the EWA unlike the IBL or the IML has a set of free parameters that must be assigned a value. These four free parameters are $\{\lambda, \rho, \delta, \phi\}$. Since our focus is on finding a learning algorithm that will be valid across games, I used a grid search to determine these values ($\lambda = 6, \rho = 0.40, \delta = 0.55, \phi = 1.00$). See Appendix A.2 for the details of the grid search.

A.2 Fitting Parameter Values

I found parameter values for both the IEL Alt and the EWA by using a rough grid search to find parameters that would minimize the mean squared error between the 500 simulated outcomes and the average frequency of entry by individuals in the last 10 rounds as reported by Duffy and Hopkins (2005) and shown in Table A.3. The focus on the last 10 rounds was an attempt to teach the algorithm to learn to reach the long run predicted outcome, which was not achievable using all rounds, the last 50 periods, or some convex combination of the three. Thus to calculate the summed mean square error, the frequency of attendance is ordered from the least likely to play *Enter* to the most likely to play *Enter* by the simulated agents. This ordered set is then compared to the experimental outcomes arranged in the same way (the last section of A.3 shows the outcomes). Let o_k^r be the k^{th} element of the ordered set for $k = 1, \ldots, 6$ for repetition r of the simulated outcomes and o_k^e be the k^{th} element of the ordered set for $k = 1, \ldots, 6$ from the experimental data. The summed mean squared error is $(\sum_{r=1}^{500} \sum_{k=1}^{6} (o_k^r - o_k^e)^2)/500$. The reason for utilizing a rather rough search instead of a much finer search or fitting via parametric analysis (at least for EWA) was an attempt to find a rough fit that could apply across multiple games. The IEL Alt appears to have some success in this area while the EWA fails to closely match the individual data.

For the IEL Alt, I limited $J = \{3, 4, 5, 6, 7\}$ to ensure a survival rate of at least 10% for the strategy with the highest forgone utility and used a grid between 0.05 and 0.95 in increments of 0.05 to determine the values of ρ and σ that minimized the sum mean squared error.

Parameters ($J = 7, \rho = 0.50, \sigma = 0.15$) provided the closest match of the data. While the model constantly introduces new strategies, the outcome produced closely matches the algorithm learning to play the PSNE(an outcome where there is no change in strategy).

For EWA, I searched for the minimum sum of mean squared error by using a grid between 0.05 and 0.95 in increments of 0.05 for ρ, a grid between 0.05 and 1.00 in increments of 0.05 for ϕ, a grid between 0.00 and 1.00 in increments of 0.05 for δ, and a grid between 0.5 and 9.5 in increments of 0.5 for λ. Perhaps the grid was too coarse for our goal of having the EWA reach the PSNE, although this may not be surprising given the failure to match play by both the Reinforcement Model and Fictitious Play model reported in Duffy and Hopkins (2005). Additionally, the initial attractions were assumed to have equal weight. Although the initial attractions have been adjusted to account for different behavior, the inclusion of more than two actions in some of the games does not appear easily transferable to a different setting.

A.3 Correcting the Type I Error Rate for Joint Hypothesis Testing

The symmetric mixed strategy Nash equilibrium produces a number of different predictions about how agents will play. The two main predictions of play expected by the algorithms are the distribution of outcomes over time and the number of agents changing their strategy per period. Thus the following hypothesis is tested:

H_0: The algorithm matches the MSNE prediction(s).

H_A: The algorithm does not.

Each hypothesis is tested with the predicted MSNE distribution relying on the application of Pearson's χ^2 Goodness of Fit test ($H_0 : (O_1, \ldots, O_k) = E(X_a, \ldots, X_k) := (p_1 * n, \ldots, p_k * n)$ where O is the observed outcome and X is the expected outcome created by the MSNE) and the per period change of actions relying on the application of a Z-test ($H_0 : \mu = \mu_0$) to the changes per period since the mean and standard deviation are easily calculated. Testing each individually is not sufficient to reject the hypothesis that the artificial agents are not playing the MSNE. Instead, both the χ^2 and the Z-test should be used simultaneously to determine whether an algorithm is producing an outcome statistically different from the MSNE[2]. In other words, there exist multiple hypotheses to test. Using the same p-value $= \alpha$ for the multiple hypotheses is unlikely to produce the correct test statistic since the probability of rejecting a true hypothesis is now $1 - (1 - \alpha)^2 = 0.0975$ when $\alpha = 5\%$. The Bonferroni correction, using $\alpha/2$, could be applied since the adjustment only involves two hypotheses, but the correction has been shown to be problematic when there is a significant correlation between the variables of interest. Column 3 of table A.1 reports the correlation between the total number of agents choosing an action[3] and the average per period changes

[2]Another alternative would be to use a series of Z-test the outcomes and expected number of agents changing their strategy per period

[3]This values is used instead of the distribution because the correlation between the choice of action is 1 for most of the games. Rapoport $2_{RC=3}$ had a correlation between the 3 actions that was not quite 1 and may explain lack of correlation, but high rejection rate for the uncorrected joint hypothesis test.

by all agents across the 10,000 iterations of the MSNE, i.e. the variable being tested in each test. Since a number of the correlations are significantly different from zero, the Bonferroni correction is likely insufficient to correctly adjust the p-value.

Table A.1: Joint Hypothesis Correction

Game	α	Correlation	Joint Test (No Correction)
D & H	0.0150	0.80^{***}	12.7%
C & P	0.0240	-0.02^{*}	9.7%
Selten	0.0210	-0.62^{***}	11.1%
Rapoport $1_{RC=2}$	0.0250	0.00	9.9%
Rapoport $2_{RC=3}$	0.0115	-0.01	19.2%
Rapoport $2_{RC=5}$	0.0250	0.00	9.5%
Erev (0.77)	0.0130	-0.78^{***}	12.6%
Erev (0.50)	0.0230	-0.01	10.7%
Erev (0.33)	0.0240	0.57^{***}	10.4%
Erev (0.22)	0.0160	0.79^{***}	11.9%
Erev (0.14)	0.0120	0.90^{***}	14.3%

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

To determine the correct p-value $= \alpha$, the outcome has been simulated with each player utilizing the MSNE in each of the games described in Section 4 and run the algorithm for 10,000 iterations. To ensure the tests being utilized were the correct test and producing the predicted outcome, I ran the Pearson's χ^2 test and the Z-test for each at $\alpha = 1\%$, 5% and 10%. Game $2_{RC=5}$ could not be directly tested via the χ^2 Goodness of Fit test because the predicted values were zero for 3 of the outcomes. Instead the game was tested as though only two of the actions were available with the tests producing the expected results. There exists no MSNE for Game $1_{RC=3}$ from Rapoport et al or the first nine games used by Erev et al. For the rest of the games table A.1 reports the adjusted α needed to recreate the 5% significance level used for a single test, the correlation between the variables of interest, and the percentage of iterations rejected without the adjustment.

The results reported for Erev et al are divided into groups based on the predicted frequency each player should play *Enter*. These probabilities match the outcomes predicted by other games allowing for a comparison across these games to be made. For instance, Duffy and Hopkins also has a predicted MSNE where the players choose to play *Enter* 22% of the time. While the *alpha* found for each game is not identical, the correlations and resulting p-value are very similar. This also appears to hold true comparing the 0.5 Erev outcome with the Rapoport game $1_{RC=2}$, game $2_{RC=5}$, and the Chmura and Pitz outcomes. There is almost no correlation between the route choice and the frequency of changing strategies with the adjustment to α also being minimal.

A.4 Additional Analysis of Games Used

This section of Appendix A gives more analysis of each individual game as well as the overlap of the games used in Chapter 2.

A.4.1 Additional Graphs, Tables, and Discussion of Duffy and Hopkins

Table A.2 gives the aggregate outcomes from applying the algorithms to the market entry game. I report the mean and the average of the standard deviation from the paper, i.e. D & H, as well as the algorithms.

Data	Attendance			Payoff		
	All Periods	Last Half	Last 10	All Periods	Last Half	Last 10
D & H	2.20	2.09	2.00	7.75	7.89	8.00[4]
	(0.72)	(0.52)	(0.21)	(0.94)	(0.58)	(0.27)
MSNE	1.32	1.32	1.32	8.00	8.00	8.00
	(1.01)	(1.01)	(0.99)	(0.56)	(0.55)	(0.48)
IBL	1.79	1.79	1.80	7.72	7.76	7.76
	(1.18)	(1.13)	(1.10)	(1.27)	(1.16)	(1.08)
IML	1.91	1.90	1.91	7.77	7.80	7.80
	(0.90)	(0.86)	(0.84)	(1.01)	(0.86)	(0.84)
EWA	1.89	1.88	1.88	7.74	7.75	7.75
	(0.11)	(0.00)	(0.00)	(0.59)	(0.49)	(0.50)
ST-EWA	2.13	2.10	2.09	7.56	7.59	7.60
	(1.11)	(1.11)	(1.08)	(1.34)	(1.31)	(1.22)
IEL	2.02	1.99	2.01	7.77	7.82	7.82
	(1.00)	(0.90)	(0.86)	(1.18)	(1.04)	(0.97)
IEL Alt	2.03	2.01	2.02	7.95	8.03	8.04
	(0.56)	(0.24)	(0.14)	(0.67)	(0.31)	(0.23)

Table A.2: Aggregate Outcomes from Algorithms and Duffy and Hopkins (2005)

Table A.3 gives the average frequency of each agent playing *Enter* when the agents are sorted (or ordered) from those who play *Enter* the least to those who play *Enter* the most. The IBL, IML, and ST-EWA all play as if each agent is learning to play some sort of mixed strategy with a slight level of heterogeneity in their actions. In other words, the agents are not learning to play how humans did in the game, but mainly on a period by period basis similar to the Logit Reinforcement Learning model described by Duffy and Hopkins. The EWA model has three players learn to stay out of the market with the other three playing in a somewhat randomized manner. Finally, the IEL and IEL Alt both come closer to representing the play displayed by the experimental subjects since 3 players learn to stay out of the market almost entirely, one player learns to always enter, and some mixing by the final two players.

[4]The value here is not based on the information reported in Table 2 (Duffy and Hopkins (2005), p. 43) since the number of agents entering the market was exactly 2 in two of the sessions and thus the average payoff should have been 8.03 instead of 8.07 reported in those two games.

Player:	1	2	3	4	5	6
Data Source	All 100 Periods					
D & H	0.01	0.09	0.19	0.44	0.67	0.79
IBL	0.25	0.27	0.29	0.30	0.32	0.36
IML	0.18	0.21	0.23	0.24	0.51	0.54
EWA	0.00	0.00	0.00	0.37	0.68	0.84
ST-EWA	0.17	0.23	0.28	0.35	0.47	0.64
IEL	0.10	0.14	0.20	0.29	0.53	0.76
IEL Alt	0.02	0.04	0.06	0.11	0.86	0.94
	Last 50 Periods					
D & H	0.00	0.03	0.07	0.23	0.79	0.97
IBL	0.22	0.26	0.28	0.31	0.34	0.38
IML	0.17	0.20	0.22	0.24	0.51	0.55
EWA	0.00	0.00	0.00	0.37	0.68	0.84
ST-EWA	0.16	0.22	0.28	0.35	0.46	0.63
IEL	0.06	0.10	0.16	0.27	0.58	0.83
IEL Alt	0.00	0.01	0.01	0.03	0.97	0.99
	Last 10 Periods					
D & H	0.00	0.00	0.00	0.10	0.90	1.00
IBL	0.13	0.21	0.27	0.32	0.39	0.48
IML	0.10	0.16	0.21	0.26	0.55	0.63
EWA	0.00	0.00	0.00	0.37	0.68	0.84
ST-EWA	0.10	0.19	0.27	0.37	0.49	0.66
IEL	0.02	0.08	0.16	0.28	0.61	0.87
IEL Alt	0.00	0.00	0.01	0.03	0.98	1.00

Table A.3: Average Frequency of Playing *Enter* Ordered from Least to Most Frequent by Player

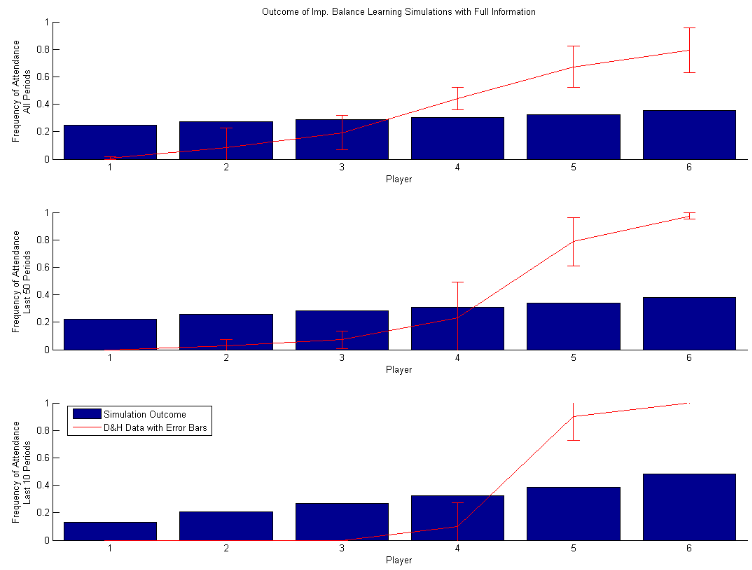


Figure A.1: Average Individual Frequency of Entering the Market with Impulse Balance Learning

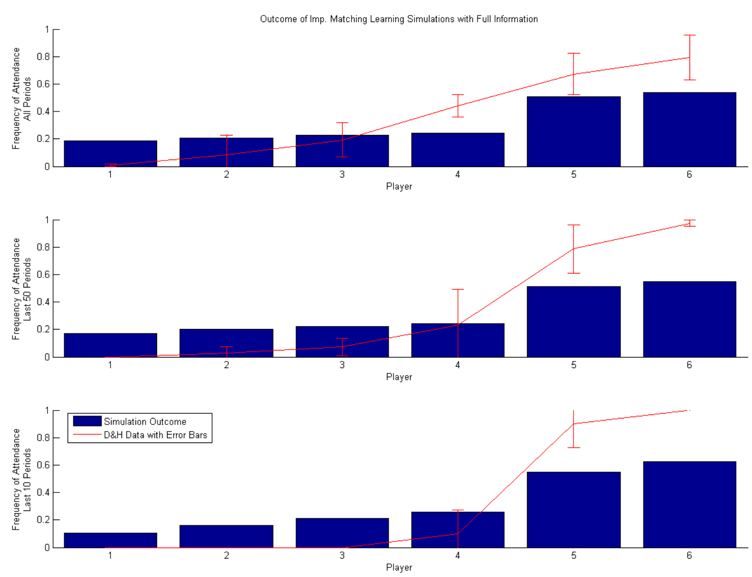


Figure A.2: Average Individual Frequency of Entering the Market with Impulse Matching Learning

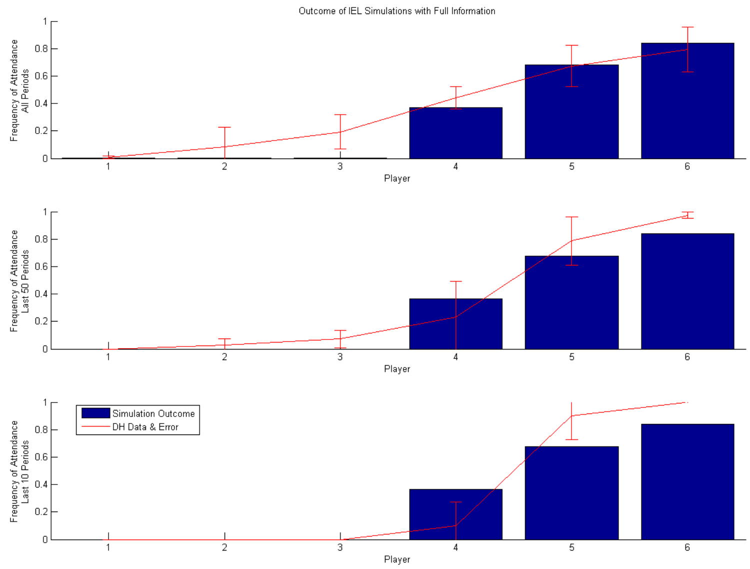


Figure A.3: Average Individual Frequency of Entering the Market with EWA Fitted Values

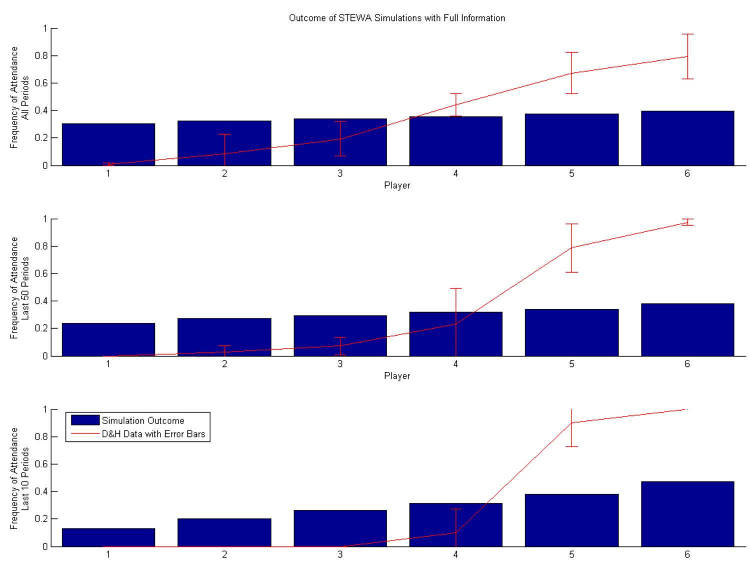


Figure A.4: Average Individual Frequency of Entering the Market with Self-Tuning EWA

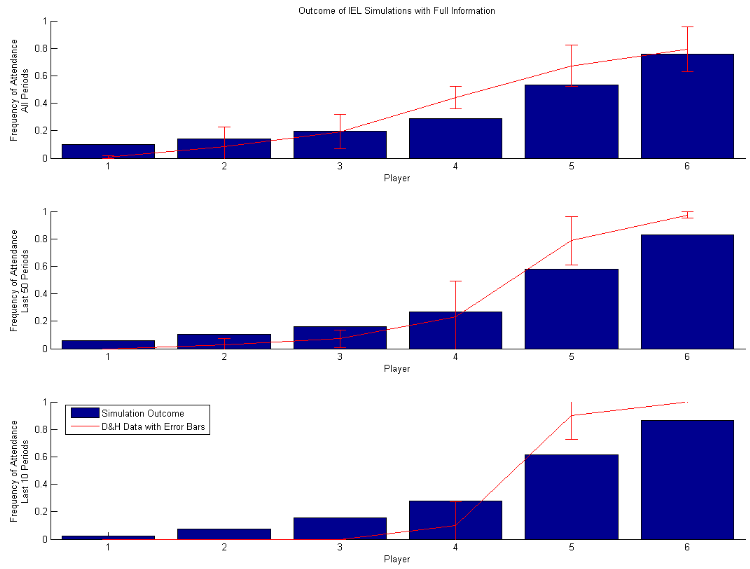


Figure A.5: Average Individual Frequency of Entering the Market with IEL

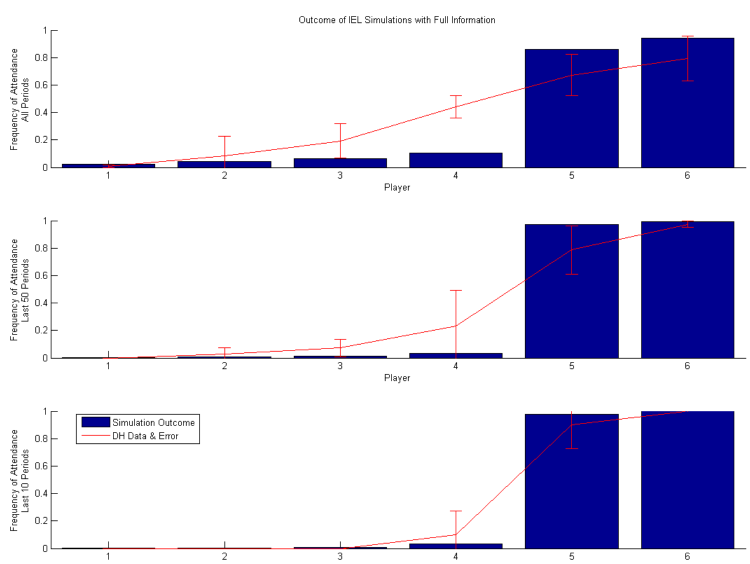


Figure A.6: Average Individual Frequency of Entering the Market with IEL Alt

A.4.2 Additional Tables and Discussion of Chmura and Pitz

Table A.4 reports the results from running the algorithms in the game used by Chmura and Pitz. The outcome from all agents playing the mixed strategy Nash equilibrium (MSNE) of the game is also reported. The EWA is the only algorithm able to achieve a higher average cumulative payoff than in the experiment. This is a result of 4 agents learning to always play Route A, 4 agents learning to always play Route B, and the final agent generally switches between the two. This behavior differs significantly from that of human subjects, but yields a cumulative payoff closer to the experimental data than the other algorithms. The MSNE and the ST-EWA are nearly indistinguishable in the outcomes produced suggesting that the ST-EWA merely replicates the mixed strategy with the selected parameter value in this setting. IBL and IML also appear to produce results similar to the MSNE. The IEL and IEL Alt both produce payoffs similar to the MSNE; but the number of road changes, the average Yule coefficient Q, and Spearman rank correlation differ from the behavior observed using the MSNE [5].

Data	Cumulative Payoff	Route A Mean	Route A Std. Dev.	Mean # of Road Changes	Yule Coef. Q Mean	Yule Coef. Q Std. Dev.	Spearman Corr. Mean	Spearman Corr. % < 0
C & P	40.76	4.44	1.01	4.27	0.56[6]	0.52	-0.54	100.0%
MSNE	36.32	4.50	1.50	4.50	0.00	0.21	-0.00	49.9%
IBL	35.83	4.50	1.56	4.64	0.05	0.20	-0.02	52.7%
IML	35.58	4.51	1.60	4.53	0.06	0.20	-0.05	54.3%
EWA	43.69	4.49	0.60	0.48	0.42	0.53	-0.68	99.2%
ST-EWA	36.29	4.50	1.50	4.50	0.01	0.20	-0.02	51.5%
IEL	36.74	4.50	1.47	3.66	0.32	0.24	-0.48	94.1%
IEL Alt	35.89	4.50	1.57	3.75	0.35	0.20	-0.45	92.7%

Table A.4: Aggregate Outcomes from the Algorithms and Chmura and Pitz (2006)

A.4.3 Additional Graphs, Tables, and Discussion of Selten et al

All of the algorithms produce results that appear to differ significantly from the behavior of subjects in the experiments as shown by Table A.5. Since the paper included a Reinforcement Learning model based on the classification of possible behavior in Table 2.6, I have included their outcome. The IBL and IML outcomes match the Pareto optimum outcome of 7 choosing route S and 11 choosing route M[7]. These also provide the Yule coefficients closest to the average behavior in the experiment. Both the EWA and ST-EWA come closer to matching the experimental data than the IBL or IML, but still fail to capture the frequency of road changes based on the RMSE. The IEL comes closest to matching the average choice of route S in the experiment, but vastly over estimates the frequency of road changes and fails to capture the response mode since the average Yule is less than zero. The IEL Alt does not appear to perform any better. Both the IEL and IEL Alt provide a better fit

[5] Table 2.5 in Appendix A.3 shows the root mean squared error produced using the means above. I continue to include the MSNE as a reference point, but do not include the table here since there is no additional insight beyond the discussion already presented.

[6] The mean for the Yule Coefficient Q and its standard deviation are taken from Chmura and Pitz (2004).

[7] See p. 396 of Selten et al (2007) for a discussion of possible equilibria in the game.

based on the RMSE value for the mean number of road changes than the other models. All of the algorithms except of the EWA maintain persistence in road changing behavior as illustrated in Figure A.7.

Data	Route S			# of Road Δ			Yule Coef. Q		
	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
Selten et al	6.06	1.67		4.27	2.16		0.28	0.60	
RL[8]	5.88	1.65	X	5.17	X	X	0.14	0.60	X
IBL	7.12	2.11	1.09	8.74	2.19	4.47	0.20	0.13	0.09
IML	6.91	2.21	0.85	8.58	2.16	4.31	0.19	0.13	0.10
EWA	6.21	0.65	0.41	0.26	1.17	4.01	0.54	0.51	0.36
ST-EWA	6.42	2.04	0.37	7.59	2.11	3.33	0.15	0.16	0.13
IEL	6.07	1.82	0.12	5.96	2.14	1.75	-0.02	0.16	0.30
IEL Alt	6.22	2.18	0.18	6.75	2.05	2.49	0.01	0.15	0.27

Table A.5: Outcomes Comparing Algorithms to Selten et al

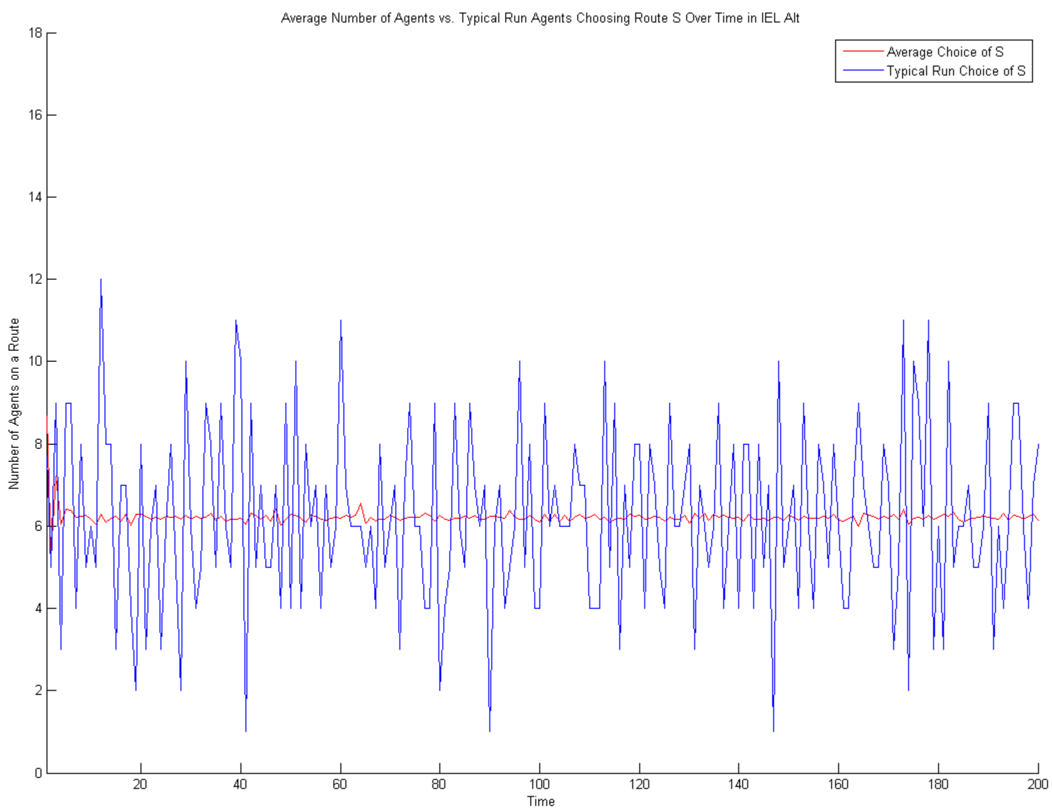


Figure A.7: Average and Typical Choice of Route S by the IEL Alt

[8]The reinforcement learning model (RL) data comes from Selten et al. The standard deviation for the number of road changes from the simulations is not reported in the paper. Because I do not have the data for the Reinforcement Learning model, I cannot include the RMSE information.

A.4.4 Additional Graphs, Tables, and Discussion of Rapoport et al

Tables A.6 and A.7 reports the average outcomes achieved by each of the algorithms with Table A.8 showing the PMSE when the costless connection is removed. All of the algorithms at the aggregate level produce nearly identical results in terms of the average number of players choosing route $O-A-D$ and $O-B-D$. With the introduction of the new node between A and B, the IBL, IML, EWA, and ST-EWA all move to the unique pure strategy Nash Equilibrium within the first couple periods. The experimental results however have a slower movement towards this pure strategy Nash Equilibrium more closely resembling the results produced by the IEL and IEL Alt.

Data	Game $1_{R=2}$		Game $1_{R=3}$		
	m_A	m_B	m_A	m_B	m_{A-B}
Rapoport et al	9.04 (2.15)	8.96 (2.15)	1.70 (1.86)	1.49 (1.68)	14.82 (2.79)
IBL	9.00 (2.53)	9.00 (2.53)	0.23 (1.03)	0.22 (1.04)	17.55 (2.02)
IML	9.00 (2.26)	9.00 (2.26)	0.15 (0.94)	0.15 (0.94)	17.70 (1.88)
EWA	9.03 (0.27)	8.97 (0.27)	0.19 (1.01)	0.19 (1.01)	17.63 (1.94)
ST-EWA	9.00 (0.23)	9.00 (0.23)	0.22 (1.04)	0.21 (1.04)	17.57 (1.98)
IEL	9.00 (2.10)	9.00 (2.10)	2.93 (1.84)	2.80 (1.92)	12.27 (2.89)
IEL Alt	9.01 (2.30)	8.99 (2.30)	1.22 (1.74)	1.25 (1.82)	15.53 (3.22)

Table A.6: Mean and Standard Deviation of Route Choice in the ADD version of Game $1_{R=2}$ and Game $1_{R=3}$

The average attendance and standard deviation achieved by the algorithms in this game are found in Table A.9. While all of the algorithms perform relatively well in replicating the outcome of game $2_{R=3}$, several of the games break down and produce results inconsistent with the experimental outcomes in game $2_{R=5}$. The IEL breaks down when there are 5 routes as shown by the different dynamics shown in Figure A.9. The difficulty for the IEL is that even though there is a probability associated with each action, the forgone utility is calculated based on what action would have been played in that period. This allows strategies to remain in A^i that may also have a high probability of playing what would have been the worst option, but simply was not selected.

A.4.5 Additional Tables and Discussion of Erev et al

The normalized mean squared deviations[9] are reported in Table A.10. This matches the analysis done in the original paper. The data used for comparison are the frequency of

[9] See Table 4 in Erev et al (2010b), p. 224.

Data	Game $1_{R=2}$		Game $1_{R=3}$		
	m_A	m_B	m_A	m_B	m_{A-B}
Rapoport et al	9.08 (2.08)	8.92 (2.08)	1.75 (1.51)	1.45 (1.49)	14.82 (2.54)
IBL	9.01 (2.71)	8.99 (2.71)	0.23 (1.04)	0.22 (1.04)	17.55 (2.03)
IML	9.00 (2.38)	9.00 (2.38)	0.31 (1.04)	0.31 (1.04)	17.37 (1.94)
EWA	9.02 (0.28)	8.98 (0.28)	0.19 (1.02)	0.19 (1.00)	17.63 (1.95)
ST-EWA	8.98 (0.22)	9.02 (0.22)	0.22 (1.05)	0.22 (1.04)	17.57 (1.99)
IEL	9.00 (2.09)	9.00 (2.09)	2.93 (1.89)	2.75 (1.91)	12.32 (2.94)
IEL Alt	9.01 (2.30)	8.99 (2.30)	1.33 (1.84)	1.32 (1.83)	15.35 (3.31)

Table A.7: Mean and Standard Deviation of Route Choice in the DELETE version of Game $1_{R=2}$ and Game $1_{R=3}$

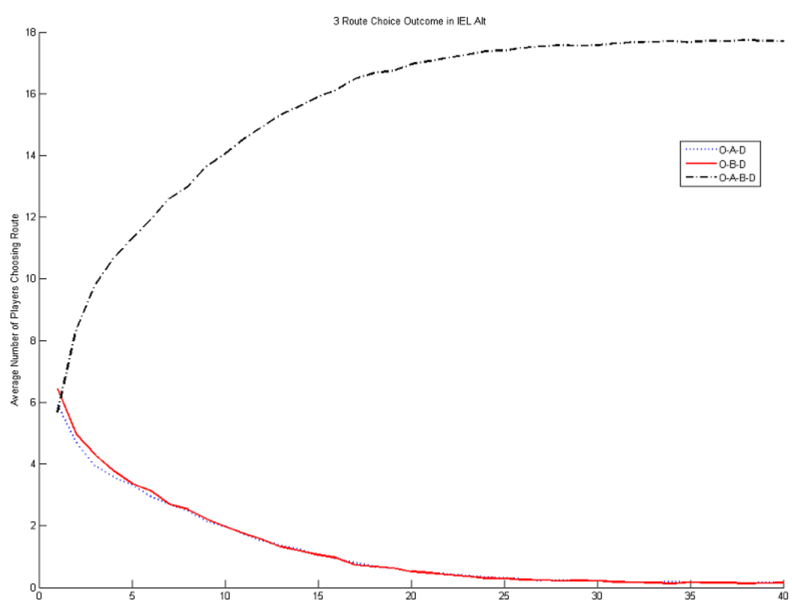


Figure A.8: Comparing Dynamics of Game $1_{RC=3}$

entry into the market (i.e. Entry), a measure of Efficiency, and the frequency that the experimental agents change to the other strategy (or Alteration). The measure of Efficiency is a simple calculation of the payoff the agents would have achieved had there been no environmental uncertainty. The data is divided into two groupings dividing the information into the first 25 rounds and second 25 rounds of the experiment, B1 and B2 respectively.

Data	Game $1_{R=2}$ RMSE m_A	Road Δ	RMSE Road Δ
IBL	0.15	20.37	7.80
IML	0.13	16.75	8.26
EWA	1.22	0.21	12.42
ST-EWA	1.15	0.15	12.46
IEL	0.28	15.96	3.54
IEL Alt	0.22	15.91	3.45

Data	Game $1_{R=3}$				
	RMSE m_A	RMSE m_B	RMSE m_{A-B}	Road Δ	RMSE Road Δ
IBL	1.53	1.23	2.73	1.33	8.49
IML	1.60	1.30	2.88	0.67	9.15
EWA	1.56	1.26	2.81	0.68	9.14
ST-EWA	1.53	1.23	2.75	0.68	9.14
IEL	1.21	1.33	2.54	17.17	7.44
IEL Alt	0.47	0.24	0.63	7.59	2.43

Table A.8: RMSE and Mean Road Changes per Agent in the DELETE version of Game $1_{R=2}$ and Game $1_{R=3}$

Data	Game $2_{R=3}$			Game $2_{R=5}$				
	m_A	m_B	m_{C-E}	m_A	m_B	m_{C-E}	m_{C-A}	m_{B-E}
Rapoport et al	5.95 (1.86)	6.01 (1.83)	6.04 (1.83)	2.08 (1.29)	1.86 (1.89)	2.06 (1.23)	5.88 (1.88)	6.13 (1.85)
IBL_i	6.02 (2.28)	6.00 (2.26)	5.98 (2.28)	0.20 (0.60)	0.18 (0.56)	0.17 (0.59)	8.72 (2.53)	8.74 (2.53)
IBL_{ii}	5.99 (2.14)	6.00 (2.14)	6.01 (2.14)	3.03 (1.64)	3.04 (1.64)	2.95 (1.62)	4.49 (1.94)	4.48 (1.94)
IML_i	6.01 (2.22)	5.85 (2.17)	6.14 (2.26)	1.08 (1.11)	1.03 (1.07)	0.99 (1.07)	7.36 (2.44)	7.54 (2.44)
IML_{ii}	6.00 (2.07)	6.00 (2.07)	6.00 (2.07)	2.97 (1.64)	2.97 (1.64)	2.89 (1.62)	4.59 (2.00)	4.59 (2.00)
EWA	6.01 (0.67)	6.00 (0.68)	5.99 (0.68)	1.43 (0.85)	1.45 (0.86)	1.50 (0.85)	6.82 (1.19)	6.80 (1.20)
ST-EWA	6.02 (0.38)	6.00 (0.38)	5.98 (0.38)	2.34 (0.56)	2.32 (0.55)	2.42 (0.50)	5.45 (0.81)	5.47 (0.80)
IEL	6.00 (1.91)	6.00 (1.93)	6.00 (1.92)	3.17 (1.58)	3.15 (1.59)	3.64 (1.68)	4.04 (1.75)	4.00 (1.72)
IEL Alt	6.00 (1.96)	5.99 (1.96)	6.01 (1.96)	2.11 (1.42)	2.11 (1.42)	2.36 (1.45)	5.70 (2.07)	5.71 (2.06)

Table A.9: Mean and Standard Deviation of Route Choice in the ADD version of Game $2_{R=3}$ and Game $2_{R=5}$

The experimental data and the measure needed to normalize the squared deviation are available via https://sites.google.com/site/gpredcomp/.

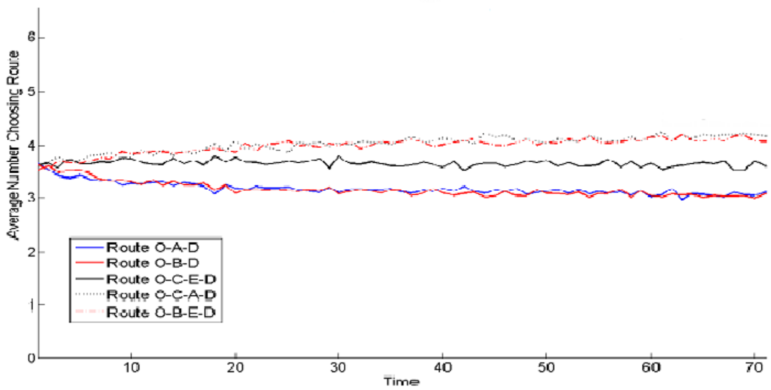


Figure A.9: Comparison of Dynamics

Model	Parameters	Entry B1	Entry B2	Efficiency B1	Efficiency B2	Alteration B1	Alteration B2	Mean
IBL	$None$	4.77	10.92	3.90	6.44	39.00	46.57	18.60
IML	$None$	4.23	6.77	2.90	4.26	17.53	34.97	11.78
EWA	$\lambda = 6.0, \phi = 1.0$ $\delta = 0.55, \rho = 0.4$	18.98	19.80	6.83	8.41	15.36	19.96	14.89
ST-EWA	$\lambda = 0.2775$	5.15	7.81	4.38	8.37	36.64	38.46	16.80
IEL	$J = 180, \rho = 0.033$ $\sigma = 0.10$	7.06	18.40	3.35	11.95	24.86	8.40	12.32
IEL Alt	$J = 7, \rho = 0.5$ $\sigma = 0.15$	13.21	25.68	5.03	13.60	5.77	8.05	11.89

Table A.10: Normalized Mean Squared Deviations

All of the algorithms perform relatively poorly compared to the algorithm that won the competition discussed in Erev et al (2010a). The winning algorithm had a mean value of the normalized squared deviations totaling ≈ 1.24, i.e. the last column of Table A.10. This is significantly lower than the results achieved by any of the algorithms implemented in the chapter. However, all of the algorithms being investigated achieve a closer fit than either the Pure Strategy Nash Equilibrium with a mean of 36.74 or the Symmetric Mixed Strategy Nash Equilibrium with a mean of 24.26. The outcome suggests that the algorithms are better approximations of human behavior than either of those two concepts for this game. This is consistent with a point Chmura et al (2012) demonstrates in a different experimental setting. The IML, IEL, and IEL Alt all achieve the lowest mean values and have an outcome approximately equivalent to the Reinforcement Learning model fit to the data in Erev et al (2010b). I do offer the caveat that none of the models implemented here are fit to this data (while the other algorithms were).

A.4.6 Analysis of Learning Behavior Relative to Experimental Data Across Games

This Appendix provides a more detailed analysis of how the algorithms perform relative to the hypothesis set forth in Chapter 2 Section 2 of this thesis.

Games with Symmetric Payoffs

Table A.11 reports the percentage of runs where the MSNE hypotheses were not rejected or the play matched the PSNE when the payoff functions were symmetric. The outcomes are grouped by the game that the algorithms played. Based on the information reported in the papers, a number of predictions are made regarding the behavior of the algorithms if they were mimicking the play of experimental subjects in these games. The first prediction is the distribution produced by the algorithms will not be statistically different from the MSNE predicted distribution using Pearson's χ^2 test shown in column 1. The next expectation is the artificial agents will not match the frequency of changes predicted by the MSNE using the Z-test shown in column 2. The Z-test alone is insufficient to reject the hypothesis that the player was using the MSNE, so the joint hypothesis test is run using the corrected p-values described in Appendix A.3 and results are given in column 3. The final prediction is all of the algorithms will fail to learn the PSNE since the human subjects participating in the experiments failed to learn to play the PSNE shown in column 4. Therefore in Table A.11, I expect column 1 to have values of 95% or higher while I expect columns 2 to 4 to report values close to 0%.

Chumra and Pitz had a unique payoff function with the cost only affecting the action taken by the majority of users resulting in outcomes indistinguishable from the distribution produced by the MSNE. The IML and ST-EWA fail to reject the joint hypothesis at nearly the same rate as they fail to reject the χ^2 test. I am able to reject the joint hypothesis for most of the runs of the IEL and IEL Alt. None of the algorithms learn to play the PSNE which is not surprising given this would involve 4 players always playing in and 5 players accepting a bad outcome[10].

The outcome reported by Rapoport et al for game $1_{RC=3}$ produced the exact mean and standard deviation predicted by the MSNE, but the average number of agents changing their action each period allowed the researchers to conclude the agents were not in fact playing the MSNE. Since the PSNE and MSNE produce the same expected distribution, failure to reject Pearson's χ^2 test could also be the result of the algorithm playing the PSNE (a result I examine). The IML is unable to reject a single iteration relative to the MSNE distribution and fails to reject nearly 80% of the runs via the joint hypothesis. Close to 19% of the iterations of the IML are consistent with the PSNE. Since the probability of half the agents choosing one route when each route is equally likely to be selected is 18.6% and the IML produced the same percentage of PSNE across all periods, the PSNE of the IML is a result of the initial selection criteria and not learning. The ST-EWA rejects the predicted

[10] Although not reported here or elsewhere in the chapter, I should point out that the Experience Weighted Attraction model described in Appendix A.1 does produce the PSNE 96% of the time. I should further note that unlike the prediction where 5 players get a terrible payoff forever, one player appears to switch between the two choices while the other 8 players maintain play the same action every time. This allows for 8 players to receive 1/2 the maximum payoff and 1 player to get zero.

Algorithm	MSNE: Last Half			PSNE
	χ^2 Test	Z-Test	Joint Test	Last Half
Game: Chmura and Pitz				
IML	98.6%	93.5%	97.0%	0.0%
ST-EWA	97.4%	93.6%	95.2%	0.0%
IEL	98.9%	2.3%	3.2%	0.0%
IEL Alt	99.6%	10.3%	15.8%	0.0%
Game: Rapoport et al $1_{RC=2}$				
IML	100.0%	75.7%	78.6%	19.3%
ST-EWA	31.5%	0.0%	0.0%	31.5%
IEL	97.7%	1.8%	2.5%	0.0%
IEL Alt	99.5%	0.0%	0.0%	0.0%
Game: Rapoport et al $2_{RC=3}$				
IML	94.7%	88.7%	89.0%	5.4%
ST-EWA	45.2%	0.0%	0.0%	54.1%
IEL	85.2%	0.1%	0.3%	0.0%
IEL Alt	99.1%	0.0%	0.2%	0.0%

Table A.11: MSNE or PSNE in Games with Symmetric Payoff Functions

MSNE for most of the runs and the joint hypothesis for all of the runs. All of the iterations consistent with the MSNE for ST-EWA are actually the result of the algorithm learning to play the PSNE. Both the IEL and IEL Alt produce similar outcomes with nearly all of the iterations rejecting the joint hypothesis (and the Z-test), while only a few iterations were able to reject the distribution predicted by the MSNE. None of the runs from either algorithm matches the PSNE as expected.

Once again the distribution of outcomes predicted by the PSNE and the MSNE are identical in game $2_{RC=3}$, but only the mean reported in Rapoport et al matches the prediction of the MSNE. Here the χ^2 test for the IML rejects 5% of the runs as not producing the MSNE distribution, an error rate consistent with the 5% significance level, but rejects the joint hypothesis 11% of the time. The number of runs using the PSNE by the IML is once again consistent with the probability that each route will be selected by six agents when each route is equally likely to be chosen. Nearly 45% of the iterations using the ST-EWA produce a distribution statistically different from the predicted MSNE distribution with almost all of the runs where the χ^2 test is not rejected using the PSNE. The IEL and the IEL Alt both reject the joint hypothesis for over 99% of the runs and neither produces runs playing the PSNE. The IEL results in a distribution that differs from the predicted MSNE at a higher rate than the previously examined games while the IEL Alt continues to have less than 1% of the runs fail the χ^2 test.

Overall, only two of the algorithms are able to consistently match the predictions made. While very few of the iterations of the IML reject MSNE distribution, the algorithm also produces runs using the PSNE and runs that fail to reject the joint hypothesis. The ST-EWA fails to perform as predicted with the algorithm producing a large number of runs consistent with the PSNE. The IML is able to produce outcomes matching the experimental data better than the ST-EWA. The IEL and the IEL Alt both produce outcomes consistent with the regularities created by human subjects in congestion games with symmetric payoffs.

Games with Asymmetric Payoffs

Table A.12 reports the percentage of runs where the MSNE hypotheses were not rejected or the play matched the PSNE with asymmetric payoff functions. The main prediction here is that the algorithms should not produce the predicted MSNE distribution in the games with full information, i.e. the distribution the algorithm produces should reject the χ^2 Goodness of Fit test using the predicted MSNE distribution. The human subjects playing the limited information games described in Erev et. al. produced outcomes where the Goodness of Fit test was not rejected for 21.8% of the experimental observations and the joint hypothesis was not rejected for 17.7% of the observations. Despite the asymmetric payoffs in Erev et al, there is no clear prediction that can be made about the MSNE hypotheses. No results are reported for the MSNE hypotheses of Rapoport game $1_{RC=3}$ because there was no MSNE for this game. There is little evidence of the algorithms using the MSNE except in the games found in Erev et al. The predictions for the PSNE are more diffuse and are discussed in more detail while reviewing the results of each paper.

The MSNE joint hypothesis test is rejected for all of the algorithms in Duffy and Hopkins. One group of human subjects, or 33% of their observations, learned to play the PSNE during the last half of the game. Only the IML and IEL Alt produce iterations where the artificial agents learn to play the PSNE. While the IML produces an outcome closer to the result reported by Duffy and Hopkins, the probability that the algorithm starts with 2 players playing enter and 4 players playing stay out is $15/64 \approx 23.4\%$ and is not far from the 24.5% outcome of the IML. The IML has 24.5% of the iterations playing this strategy for all 100 periods, implying the outcome is not the result of learning. Some of the runs of the IEL Alt learn to play the PSNE strategy since 0% of the iterations played this for all 100 periods.

Examining the performance of the algorithms in the route choice game used by Selten et al, I am able to reject the joint hypothesis for all iterations of both the IML and ST-EWA. The joint hypothesis is not rejected for all iterations of the IEL and IEL Alt, but over 99% of the runs reject the joint hypothesis. Over $2/3^{rd}$s of runs the IEL produces are statistically indistinguishable from the MSNE distribution while $1/3^{rd}$ of the experimental observations failed to reject the MSNE distribution for the same p-value. The IEL Alt fails to reject 1/5 of the runs providing an outcome closer to the experimental data. Additionally, none of the algorithms learns to play the PSNE. The lack of support for the algorithm playing the MSNE or mimicking the PSNE is consistent with the data produced by human subjects. The final issue is to determine if the aggregate distribution of play produced by each algorithm is similar to the distribution of play predicted by the PSNE[11]. I conducted another χ^2 Goodness of Fit Test using the distribution predicted by the PSNE to compare the aggregate distribution produced by the algorithms. I am able to reject the null hypothesis that the aggregate outcome created by the IML and the ST-EWA is consistent with the PSNE at any reasonable p-value, but I am not able to reject the null for the IEL or the IEL Alt at a 10%, 5%, or 1% significance level.

[11]The aggregate outcome of the experimental data is statistically indistinguishable from the PSNE.

The two games with asymmetric payoffs used in Rapoport et al are examined together. Game $1_{RC=3}$ did not have a MSNE, so the algorithms are examined to determine if the PSNE is learned. The human subjects were approaching the PSNE, but never played the PSNE for any sustained time frame in Game $1_{RC=3}$. Both the IML and the ST-EWA start playing the PSNE in the 2^{nd} period and maintain the same actions for the remainder of the game. All of the runs of the IEL fail to learn the PSNE and only a minor fraction of the IEL Alt learn to play the PSNE. I am able to reject the null hypothesis that the computational agents are using the MSNE distribution or adjusting their strategies as predicted by the MSNE for the IML, IEL, and IEL Alt for game $2_{RC=5}$, but this is not the case for a small number of iterations of the ST-EWA. The joint hypothesis is rejected for all of the algorithms. None of the algorithms learn to play the PSNE, just as none of the experimental outcomes produced the PSNE.

Games 1-9 used by Erev et al did not have a MSNE, thus none of the hypotheses are tested for these games and they are not included in the averages. The PSNE gives the average across all 40 games. The IML creates an entry rate that closely matches the data

Algorithm	MSNE: Last Half			PSNE
	χ^2 Test	Z-Test	Joint Test	Last Half
Game: Duffy and Hopkins				
IML	0.5%	9.9%	0.0%	24.5%
ST-EWA	0.0%	45.5%	0.0%	0.0%
IEL	0.8%	39.6%	0.0%	0.0%
IEL Alt	0.0%	0.3%	0.0%	15.1%
Game: Selten et al				
IML	0.0%	31.8%	0.0%	0.0%
ST-EWA	0.0%	83.0%	0.0%	0.0%
IEL	67.4%	1.1%	0.6%	0.0%
IEL Alt	20.0%	11.9%	0.2%	0.0%
Game: Rapoport et al $1_{RC=3}$				
IML	N/A	N/A	N/A	100.0%
ST-EWA	N/A	N/A	N/A	99.0%
IEL	N/A	N/A	N/A	0.0%
IEL Alt	N/A	N/A	N/A	0.2%
Game: Rapoport et al $2_{RC=5}$				
IML	0.0%	0.0%	0.0%	0.0%
ST-EWA	4.4%	0.0%	0.0%	0.0%
IEL	0.0%	0.0%	0.0%	0.0%
IEL Alt	0.0%	0.0%	0.0%	0.0%
Games: Erev et al				
IML	37.2%	52.5%	38.6%	1.0%
ST-EWA	28.8%	51.6%	30.0%	0.1%
IEL	23.5%	34.5%	19.3%	0.0%
IEL Alt	18.5%	14.9%	5.4%	5.1%

Table A.12: MSNE or PSNE in Games with Asymmetric Payoff Functions based on the root means square error with the ST-EWA producing approximately the same outcome and both produce the fewest rejections of the MSNE hypotheses. The IEL Alt has the highest rejection rate for the MSNE hypotheses. I am not able to reject the joint hypothesis for nearly 20% of the IEL runs a rate similar to the outcome produced when examining the experimental data. The Erev data did show approximiately 5.5% of the human subjects learning to play the PSNE in games 1-9 during the second half of the game. The percentage of iterations where the IEL Alt learns to play the PSNE is too high since the average covers all 40 games while the IEL fails to have any runs learn the PSNE, a percentage that is too low.

When there is full information, all of the algorithms do a fairly consistent job of producing outcomes that are statistically different from MSNE predictions. Only the IEL Alt consistently produces outcomes similar to the experimental data regarding the PSNE. The IEL never learns to play the PSNE with any measurable regularity in any of the games. If the IML and ST-EWA learn to play the PSNE, the algorithms learn to do so starting in the second repetition of play. While this may be learning, it is not the slow dynamic process displayed by subjects in the lab. From this I can conclude that the IEL Alt best captures the play in full information games, but is not able to replicate the experimental play seen in the limited information games.

Appendix B

Additional Material: Chapter 3

In this Appendix we report on outcomes from simulating the IEL Model using the parameter vector that was found to be a best fit to the experimental data.

Recall that for BoS, the grid search applied to the BoS Strangers Direct Recommendations treatment data yielded the best fitting parameter vector of $(J,\ \rho,\ \theta) = (190, 0.05, 0.125)$. The tables and figures that follow are the analogues of Tables 3.3-3.4, Figures 3.1-3.2, Tables 3.5-3.6, Figures 3.3, Tables 3.7-3.8, and Figures 3.4-3.5 reported on in Section 4.1 but using this best-fit parameter vector in place of the baseline parameterization.

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.911 (0.092) [0.841,0.982]	5.47 (0.548) [5.04,5.89]	0.993 (0.015) [0.982,1.00]	5.96 (0.088) [5.89,6.02]
Sim. Strangers	0.877 (0.046) [0.873,0.881]	5.26 (0.276) [5.24,5.28]	0.979 (0.040) [0.976,0.983]	5.87 (0.242) [5.85,5.90]
DLL Partners	0.857 (0.169) [0.727,0.988]	5.14 (1.016) [4.36,5.93]	0.878 (0.331) [0.624,1.132]	5.27 (1.886) [3.74,6.79]
Sim. Partners	0.862 (0.081) [0.855,0.869]	5.17 (0.487) [5.13,5.21]	0.963 (0.086) [0.956,0.971]	5.78 (0.514) [5.73,5.82]

Table B.1: BoS Direct Recommendations Comparison, Best-Fit Parameter Vector

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
DR	T = 1-60	0.0%	0.0%	0.0%	0.0%	0.85
Strangers	31-60	0.0%	0.0%	0.0%	0.0%	0.96
	46-60	0.2%	0.1%	0.0%	0.1%	0.97
DR	T = 1-60	0.0%	0.0%	0.0%	0.4%	0.79
Partners	31-60	2.8%	2.2%	5.0%	3.4%	0.88
	46-60	2.8%	2.0%	5.6%	3.2%	0.89

Table B.2: IEL nf and f Strategies in BoS-DR, Best-Fit Parameter Vector

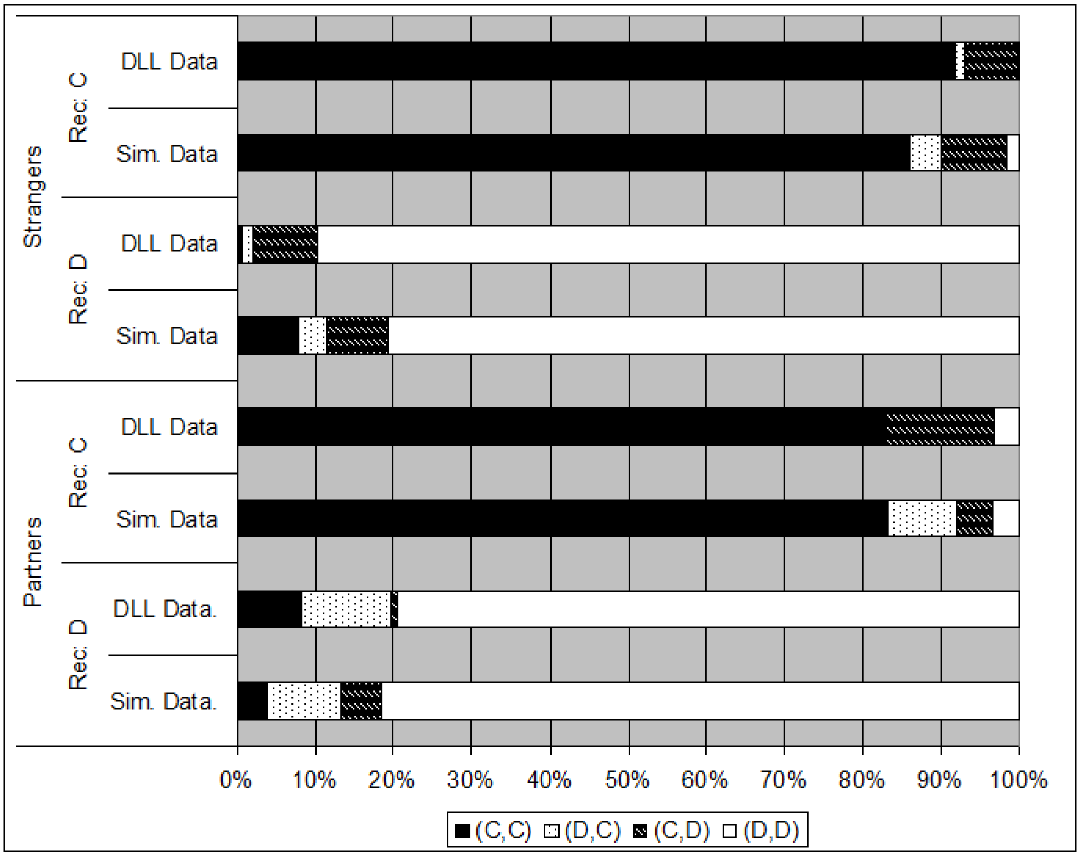
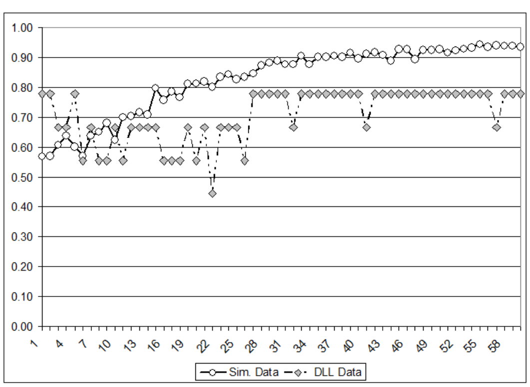
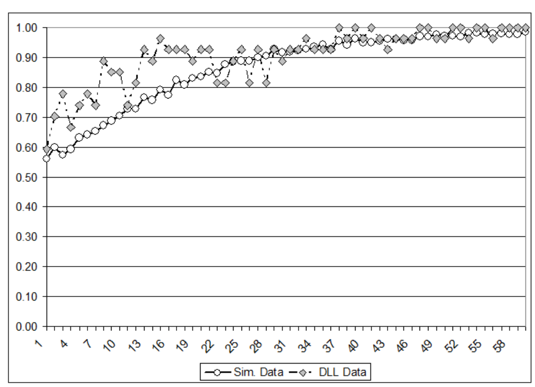


Figure B.1: Outcome Frequencies Based on Recommendations in BoS-DR, Best Fit Parameter Vector



(a) BoS-DR Partners



(b) BoS-DR Strangers

Figure B.2: Pairwise Recommendation Following over Time in the BoS-DR treatments: Simulations versus DLL Data, Best Fit Parameter Vector

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.548 (0.141) [0.439,0.656]	3.29 (0.843) [2.64,3.93]	0.596 (0.203) [0.440,0.753]	3.58 (1.218) [2.64,4.51]
Sim. Strangers	0.827 (0.076) [0.820,0.833]	4.96 (0.454) [4.92,5.00]	0.979 (0.056) [0.974,0.984]	5.87 (0.336) [5.84,5.90]
DLL Partners	0.869 (0.185) [0.726,1.011]	5.21 (1.111) [4.36,6.06]	0.800 (0.316) [0.557,1.043]	4.80 (1.897) [3.34,6.26]
Sim. Partners	0.895 (0.069) [0.889,0.901]	5.37 (0.416) [5.33,5.40]	0.974 (0.059) [0.969,0.980]	5.85 (0.352) [5.82,5.88]

Table B.3: BoS No Recommendations (None) Comparison, Best-Fit Parameter Vector

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
None	T = 1-60	0.5%	0.7%	0.3%	0.9%	0.00
Strangers	31-60	17.1%	21.7%	22.7%	21.4%	0.00
	46-60	19.7%	24.8%	25.6%	24.3%	0.00
None	T = 1-60	2.8%	3.0%	2.5%	3.2%	0.00
Partners	31-60	24.8%	26.5%	23.6%	24.4%	0.00
	46-60	24.7%	26.5%	24.0%	24.4%	0.00

Table B.4: IEL nf and f Strategies in BoS-None, Best-Fit Parameter Vector

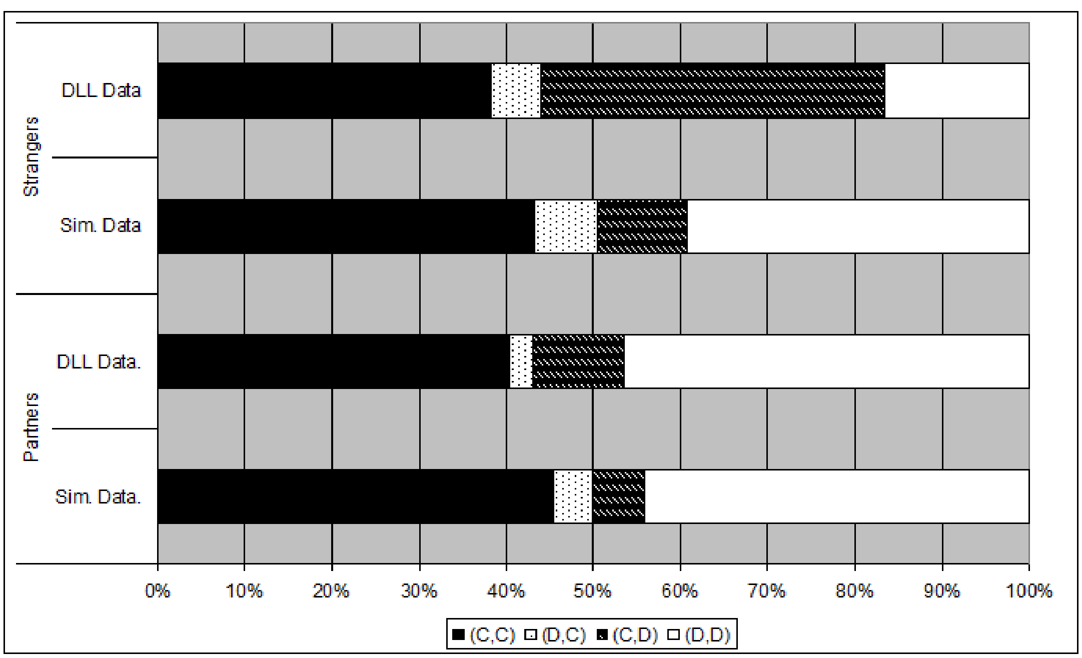


Figure B.3: Observed Outcome Frequencies in BoS-None, Best-Fit Parameter Vector

Type of Set-Up	Coordination All Rnds.	Avg. Payoff All Rnds.	Coordination Last 10	Avg. Payoff Last 10
DLL Strangers	0.580 (0.119) [0.488,0.671]	3.48 (0.715) [2.93,4.03]	0.674 (0.174) [0.532,0.816]	4.04 (1.108) [3.19,4.90]
Sim. Strangers	0.608 (0.086) [0.600,0.615]	3.65 (0.516) [3.60,3.69]	0.735 (0.162) [0.721,0.749]	4.41 (0.972) [4.32,4.49]
DLL Partners	0.915 (0.128) [0.808,1.022]	5.49 (0.768) [4.84,6.13]	1.000 (0.000) [1.000,1.000]	6.00 (0.000) [6.00,6.00]
Sim. Partners	0.730 (0.143) [0.718,0.743]	4.38 (0.856) [4.31,4.46]	0.905 (0.167) [0.891,0.920]	5.43 (1.002) [5.34,5.52]

Table B.5: BoS Indirect Recommendations Comparison, Best-Fit Parameter Vector

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
IR.	T = 1-60	0.0%	0.0%	0.0%	0.0%	0.39
Strangers	31-60	4.9%	6.1%	3.9%	4.1%	0.33
	46-60	7.8%	10.3%	6.7%	8.5%	0.30
IR.	T = 1-60	0.3%	0.0%	0.1%	0.0%	0.28
Partners	31-60	22.3%	23.7%	10.4%	11.4%	0.15
	46-60	26.9%	29.4%	12.3%	13.7%	0.11

Table B.6: IEL nf and f Strategies in BoS-IR, Best-Fit Parameter Vector

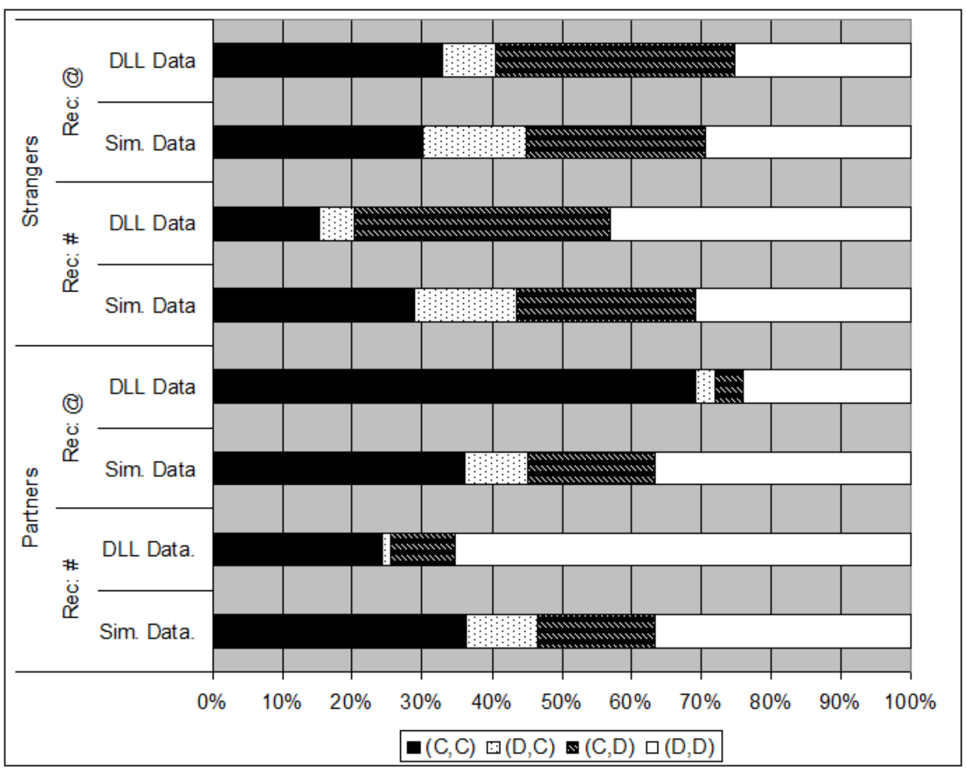
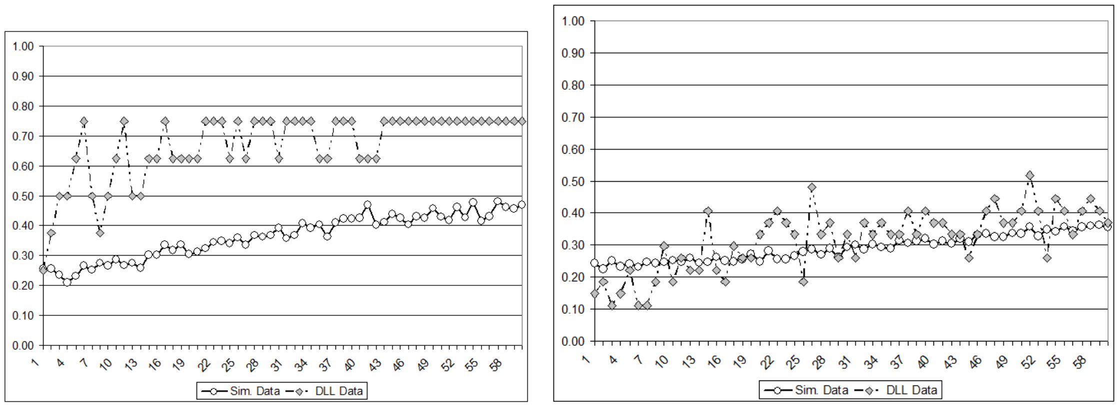


Figure B.4: Observed Outcome Frequencies in BoS-IR, Best-Fit Parameter Vector



(a) BoS-IR Partners

(b) BoS-IR Strangers

Figure B.5: Pairwise Recommendation Following over Time in the BoS-IR treatments: Simulations versus DLL Data, Best Fit Parameter Vector

Recall that for Chicken, the best fitting parameter vector was $(J,\ \rho,\ \theta) = (170, 0.25, 0.05)$. The tables and figures that follow are the analogues of Tables 3.9-3.11 and Figures 3.6-3.7 reported in Section 4.2 but using this best-fit parameter vector in place of the baseline parameterization.

Data Source:	Good	Nash	Bad	Very Good	None
DF Data	100.81	104.59	95.87	99.70	99.76
Baseline	101.76	101.40	91.77	104.53	95.82
Grid Search Fit	102.58	101.16	90.00	106.01	93.96

Table B.7: Comparison of Efficiency Measures in all Chicken Treatments, Best-Fit Parameter Vector

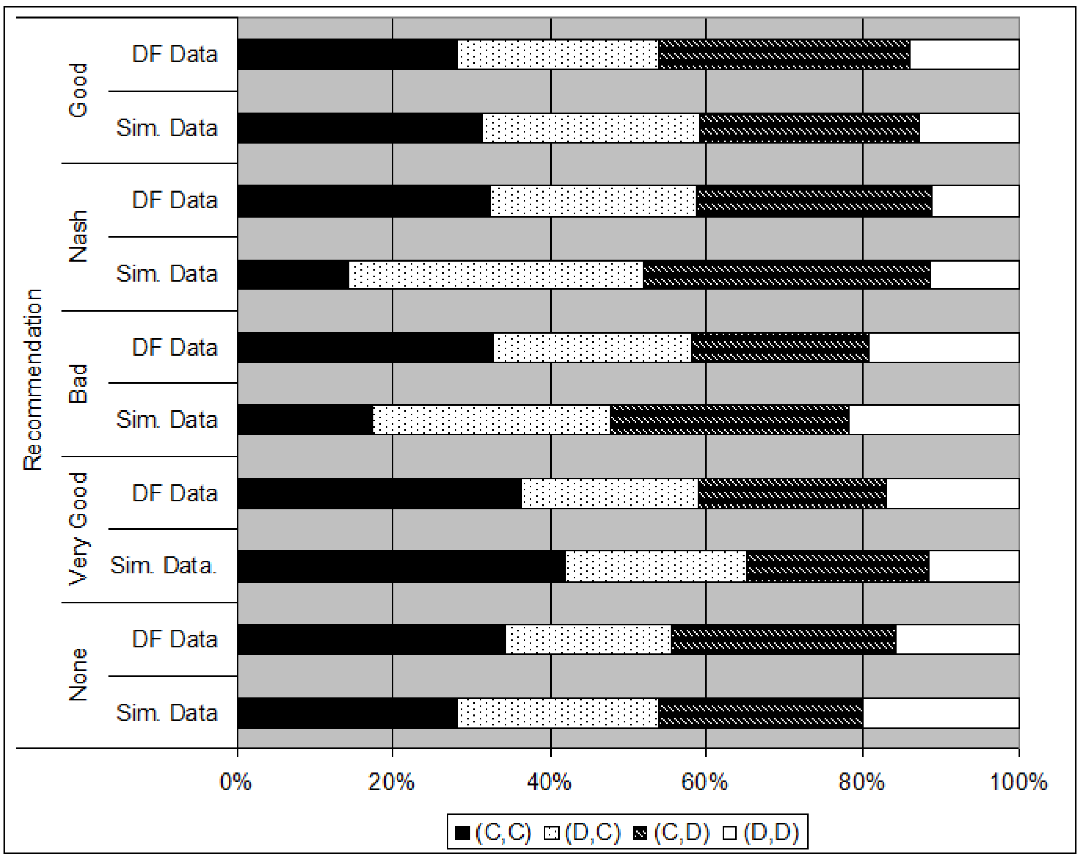


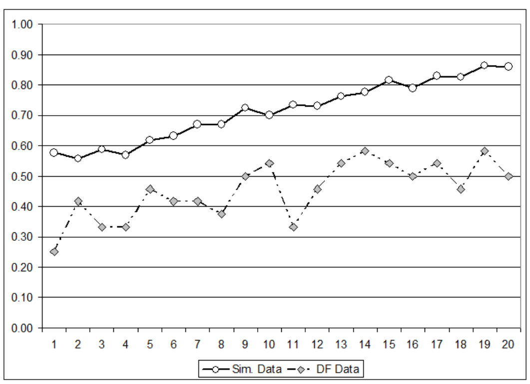
Figure B.6: Observed Outcome Frequencies in all Chicken Treatments, Best-Fit Parameter Vector

% Follow	Good Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	73.5%	78.6%	0.011	73.0%	77.5%	0.012	75.0%	82.0%	0.008
Rec = C	73.2%	78.1%	0.012	71.9%	77.4%	0.013	77.0%	80.2%	0.008
Total	73.3%	78.3%	0.009	72.3%	77.4%	0.011	76.2%	80.8%	0.004
Pairwise	53.1%	61.5%	0.022	51.4%	60.2%	0.026	58.3%	65.4%	0.010
% Follow	Nash Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	56.7%	82.8%	0.076	55.3%	80.2%	0.071	60.8%	90.5%	0.090
Rec = C	77.7%	85.7%	0.015	77.2%	83.6%	0.011	79.2%	92.0%	0.025
Total	67.2%	84.2%	0.032	66.3%	81.9%	0.027	70.0%	91.3%	0.047
Pairwise	45.4%	71.5%	0.073	43.3%	67.5%	0.063	51.7%	83.3%	0.102
% Follow	Bad Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	47.7%	75.5%	0.084	46.0%	74.4%	0.088	52.9%	78.7%	0.071
Rec = C	63.1%	82.7%	0.063	66.5%	81.6%	0.039	53.0%	85.9%	0.136
Total	54.1%	78.4%	0.063	54.5%	77.3%	0.056	52.9%	81.6%	0.086
Pairwise	26.9%	61.3%	0.125	25.9%	59.6%	0.120	30.0%	66.6%	0.138
% Follow	Very Good Recommendation								
	DF	Sim.	MSE	DF	Sim.	MSE	DF	Sim.	MSE
	T = 1-20			T = 1-15			T = 16-20		
Rec = D	51.1%	75.4%	0.120	60.6%	75.9%	0.059	22.7%	73.8%	0.289
Rec = C	60.8%	69.7%	0.014	63.2%	70.4%	0.011	53.7%	67.5%	0.021
Total	59.9%	70.2%	0.018	62.9%	70.9%	0.012	50.8%	68.2%	0.033
Pairwise	36.5%	49.0%	0.027	40.0%	50.2%	0.021	25.8%	45.7%	0.044

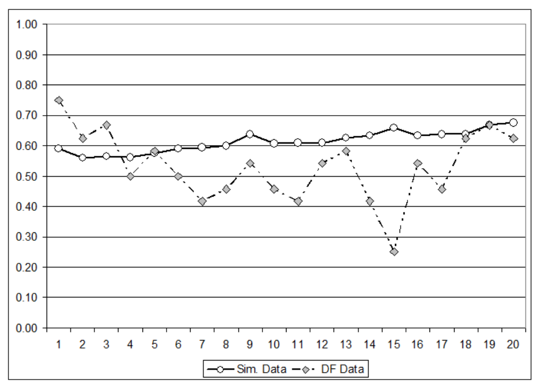
Table B.8: Observed Frequencies of Recommendation Following in all Chicken Treatments, Best-Fit Parameter Vector

Type	Time	[D D]	[C C]	[D C]	[C D]	Avg. f
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.57
Good	11-20	0.0%	0.0%	0.0%	0.0%	0.61
	16-20	0.0%	0.0%	0.0%	0.0%	0.63
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.68
Nash	11-20	0.0%	0.0%	0.0%	0.0%	0.79
	16-20	0.0%	0.0%	0.0%	0.0%	0.83
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.59
Bad	11-20	0.0%	0.0%	0.0%	0.0%	0.64
	16-20	0.0%	0.0%	0.0%	0.0%	0.66
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.45
Very Good	11-20	0.0%	0.0%	0.0%	0.0%	0.43
	16-20	0.0%	0.0%	0.0%	0.0%	0.42
Simulation	T = 1-20	0.0%	0.0%	0.0%	0.0%	0.00
None	11-20	0.0%	0.0%	0.0%	0.0%	0.00
	16-20	0.0%	0.0%	0.0%	0.0%	0.00

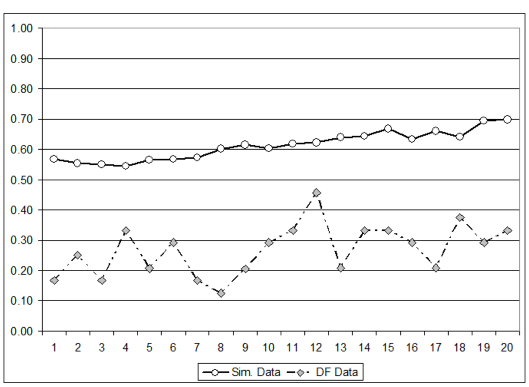
Table B.9: IEL nf and f Strategies in all Chicken Treatments, Best-Fit Parameter Vector



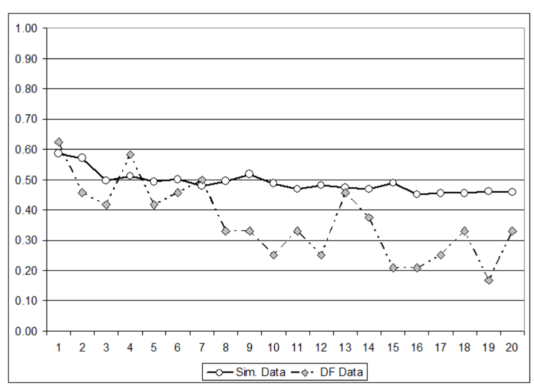
(a) Chicken-Nash



(b) Chicken-Good



(c) Chicken-Bad



(d) Chicken-Very Good

Figure B.7: Pairwise Recommendation Following over Time in the Chicken Treatments: Simulations versus DF Data, Best-Fit Parameter Vector

Appendix C

Additional Material: Chapter 4

The full formula for the various scoring rules are presented here. Let b^i be the agent's belief of the action of others, let $\bar{c}^{-i}$ be the average contribution of others, and let $\sum_{-i} c^{-i}$ be the total contribution of others. I will use b^i for the belief or prediction regardless of of whether the agents are trying to predict the average or aggregate behavior. The scoring rule used by Fischbacher & Gächter (2010) is:

$$\pi^i_{b,t}(b^i|round(\bar{c}^{-i})) = \left\{ \begin{array}{c} 3 \\ 2 \\ 1 \\ 0 \end{array} \right\} \text{if} \left\{ \begin{array}{l} |b^i - round(\bar{c}^{-i})| = 0 \\ |b^i - round(\bar{c}^{-i})| = 1 \\ |b^i - round(\bar{c}^{-i})| = 2 \\ Otherwise \end{array} \right\}.$$

Gächter & Renner (2010) implemented:

$$\pi^i_{b,t}(b^i|\bar{c}^{-i}) = \left\{ \begin{array}{c} 20 \\ \frac{10}{|b^i - \bar{c}^{-i}|} \end{array} \right\} \text{if} \left\{ \begin{array}{l} |b^i - \bar{c}^{-i}| <= 1 \\ Otherwise \end{array} \right\}.$$

The scoring rule in Croson (2000) was:

$$\pi^i_{b,t}(b^i|\sum_{-i} c^{-i}) = \left\{ \begin{array}{c} 25 \\ \frac{12.5}{|b^i - \sum_{-i} c^{-i}|} \end{array} \right\} \text{if} \left\{ \begin{array}{l} |b^i - \sum_{-i} c^{-i}| = 0 \\ Otherwise \end{array} \right\}.$$

Nugebauer et al (2009):

$$\pi^i_{b,t}(b^i|\sum_{-i} c^{-i}) = \frac{1}{400} \times (100 - |b^i - \sum_{-i} c^{-i}|)^2$$

A. Smith (2013):

$$\pi^i_{b,t}(b^i|\bar{c}^{-i}) = \left\{ \begin{array}{c} 2 - 2 \times |b^i - \bar{c}^{-i}| \\ 0 \end{array} \right\} \text{if} \left\{ \begin{array}{l} |b^i - \bar{c}^{-i}| < 1 \\ Otherwise \end{array} \right\}.$$

Chaudhuri & Paichayontvijit (2010), only elicited in the first round for the majority of the treatments. Did not elicit every period in any of the treatments.:

$$\pi^i_{b,t}(b^i|round(\bar{c}^{-i})) = \$1.00 - \$(0.1 \times (b^i - round(\bar{c}^{-i})))^2$$