Learning Correlated Equilibria: An Evolutionary Approach*

Jasmina Arifovic[†] Joshua F. Boitnott[‡] John Duffy[§]

Forthcoming in
Journal of Economic Behavior and Organization

Abstract

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 paper 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 that 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.

Keywords: Correlated equilibrium, learning, evolutionary algorithms, adaptation, game theory, experimental economics.
JEL Codes: C63, C72, C73, D83.

*We thank the editor, Nobuyuki Hanaki, two anonymous referees and audience members at the Canadian Economics Association Conference and the Workshop on the Economic Science with Heterogeneous Interacting Agents for helpful comments and suggestions on earlier drafts. We also thank Nick Feltovich, Ernest K. Lai and Wooyoung Lim for the use of their data.

[†]Department of Economics, Simon Fraser University, Email: arifovic@sfu.ca
[‡]Department of Economics, Simon Fraser University, Email: jfb@sfu.ca
[§]Department of Economics, University of California, Irvine, Email: duffy@uci.edu

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 paper 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.

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 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 1 and the Chicken game shown in Table 2.

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

Table 1: Battle of the Sexes

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

Table 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 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.

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

[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.

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:

$$S_t^i = s_{1,t}^i, s_{2,t}^i \cdots s_{J,t}^i$$

with $s_{j,t}^i = (f_{j,t}^i, nf_{j,t}^i) \forall j = 1, \ldots, J$. 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 $\ell(nf) = 1$.

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 $\rho/2$.

[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.

[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.

Computation of **Foregone Payoffs** determines the expected foregone payoff that the player could have achieved from a strategy $s^i_{j,t} \in S^i_t$ taking player i's history, h^i_t, as given. The history, h^i_t, 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:

$$W(s^i_{j,t}|h^i_t) = f^i_j \sum_{k=0}^{1} u(r^i_{t-k}, \sigma^{-i}_{t-k}) + (1 - f^i_j) \sum_{k=0}^{1} u((nf)^i_{j,t-k}, \sigma^{-i}_{t-k}) \tag{1}$$

where $W(s^i_{j,t}|h^i_t)$ is the expected foregone payoff, r^i_{t-k} is the recommendation at $t-k$, and σ^{-i}_{t-k} is the action of player $-i$ at $t-k$.

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

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 $W(s^i_{j,t}|h^i_t) = W^i_{j,t}$, two strategies $s^i_{l,t}$ and $s^i_{k,t} \in S^i_t$ are chosen randomly (with uniform probability and replacement) to compete 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]

[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.

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 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.

4.1 Battle of the Sexes

Duffy et al. (2016) (henceforth DLL) study the stage game (Table 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 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 pe-

[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 $t + 1$.

riod, 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} × {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,

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

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 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^i_t, 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^i_{t+1}, 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: BoS Direct Recommendations Comparison

Table 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 con-

[8]Later, in section 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 (and later for Tables 5 and 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.

fidence 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 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 standard deviations in the humans subject data are generally greater than in the simulated data. Perhaps the most important finding from Table 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 4: IEL nf and f Strategies in BoS-DR

Table 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 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.

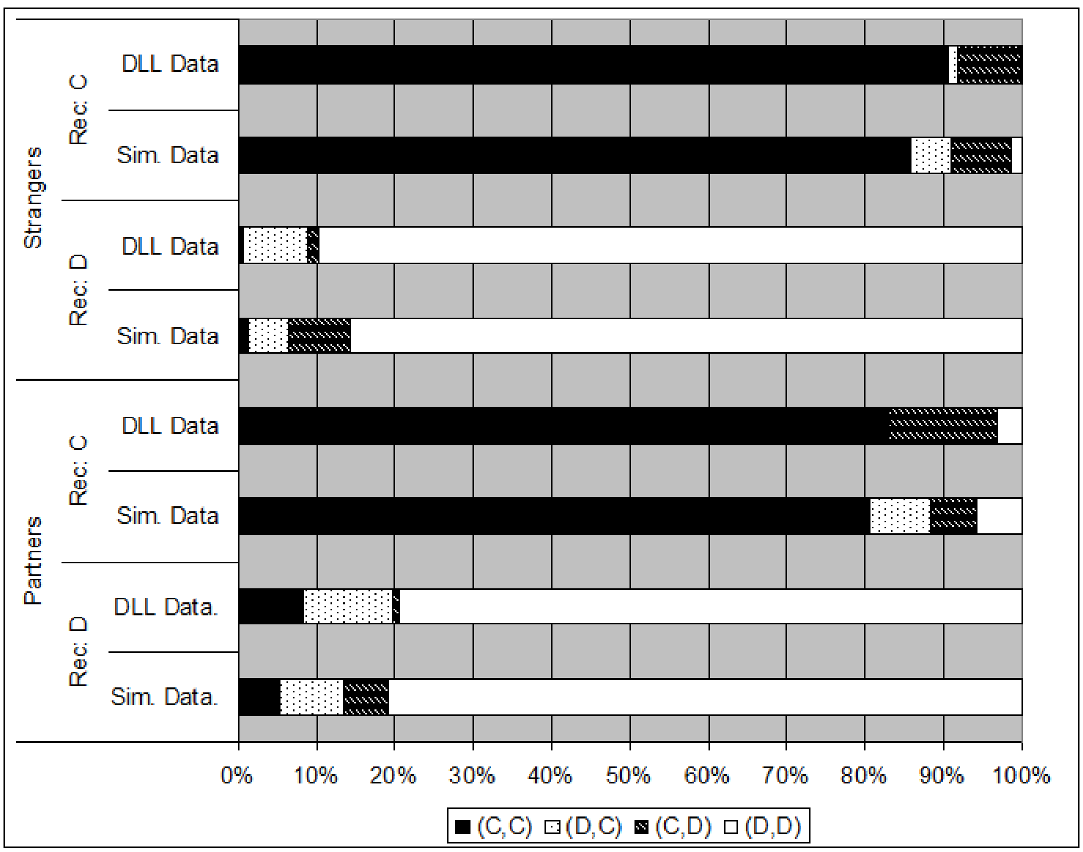
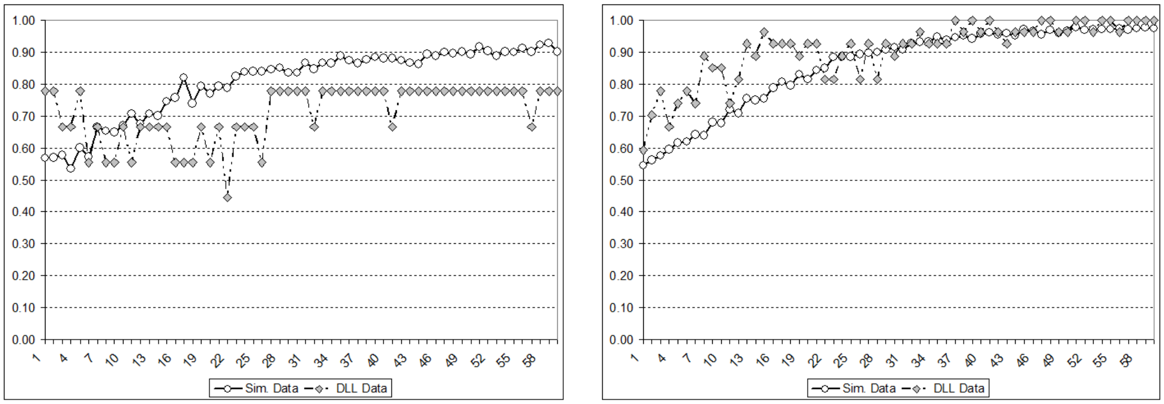


Figure 1: Outcome Frequencies Based on Recommendations in BoS-DR

Figure 1 provides a more disaggregated comparison of the mean frequencies of actual



(a) BoS-DR Partners (b) BoS-DR Strangers

Figure 2: Pairwise Recommendation Following over Time in the BoS-DR treatments: Simulations versus DLL Data

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 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 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 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 5. The simulation results for this case are obtained by setting $f^i_j = 0\, \forall j, i$ 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 6. Reported results are again from 500 simulations, and, as in Table 3, Table 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 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 (D,D) which are too low and too high (respectively) in the simulated Strangers treatment relative to comparable DLL Strangers data.[10]

Table 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 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.

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.

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 5: BoS No Recommendations (None) Comparison

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 6: IEL nf and f Strategies in BoS-None

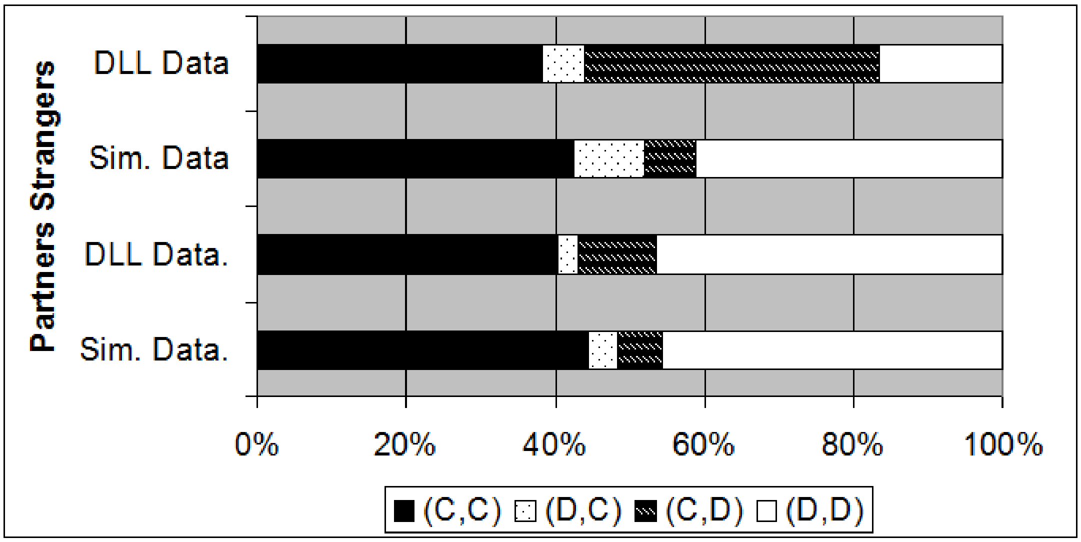


Figure 3: Observed Outcome Frequencies in BoS-None

#, 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:

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

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

Table 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 8 which uses the strategy classification discussed earlier in connection with Table 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 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 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 8: IEL nf and f Strategies in BoS-IR

Finally, Figure 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 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

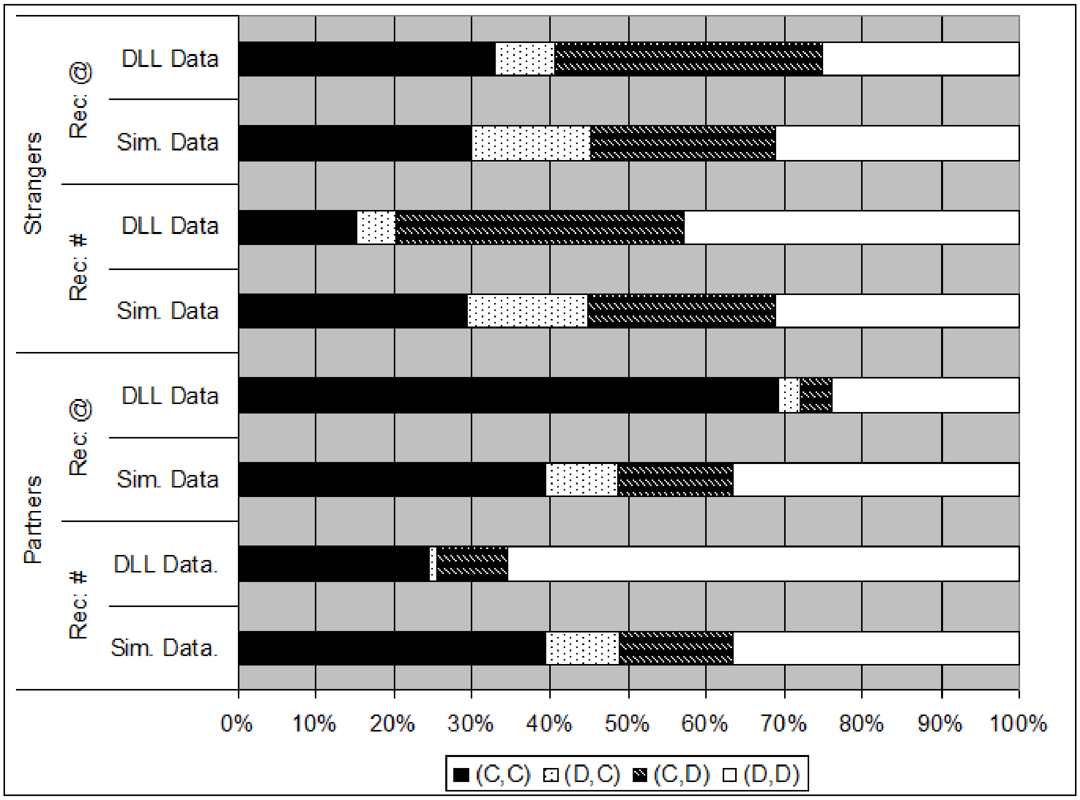


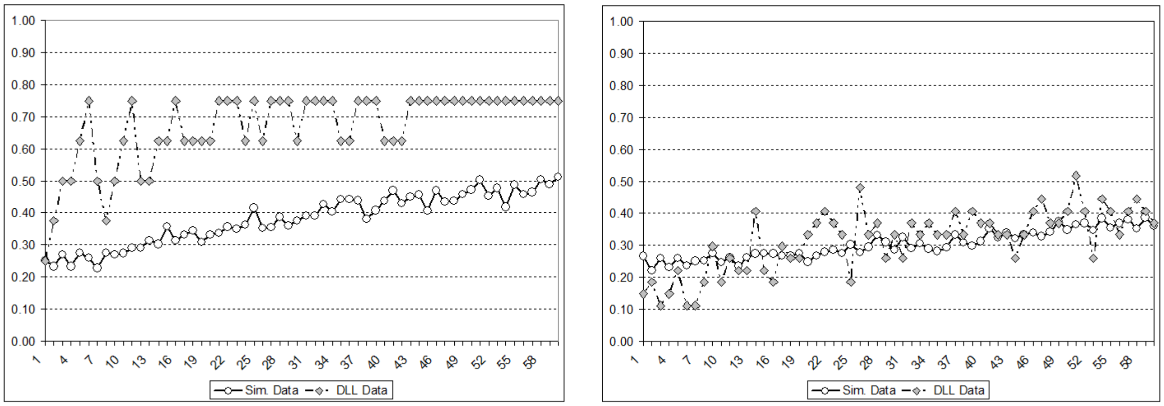
Figure 4: Outcome Frequencies based on Recommendations in BoS-IR

BoS game.

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 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



(a) BoS-IR Partners

(b) BoS-IR Strangers

Figure 5: Pairwise Recommendation Following over Time in the BoS-IR treatments: Simulations versus DLL Data

the Battle of the Sexes game, but here, if the row player is recommended to play C, the 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 9 and Figure 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 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 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

[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.

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 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 9: Comparison of Efficiency Measures in all Chicken Treatments

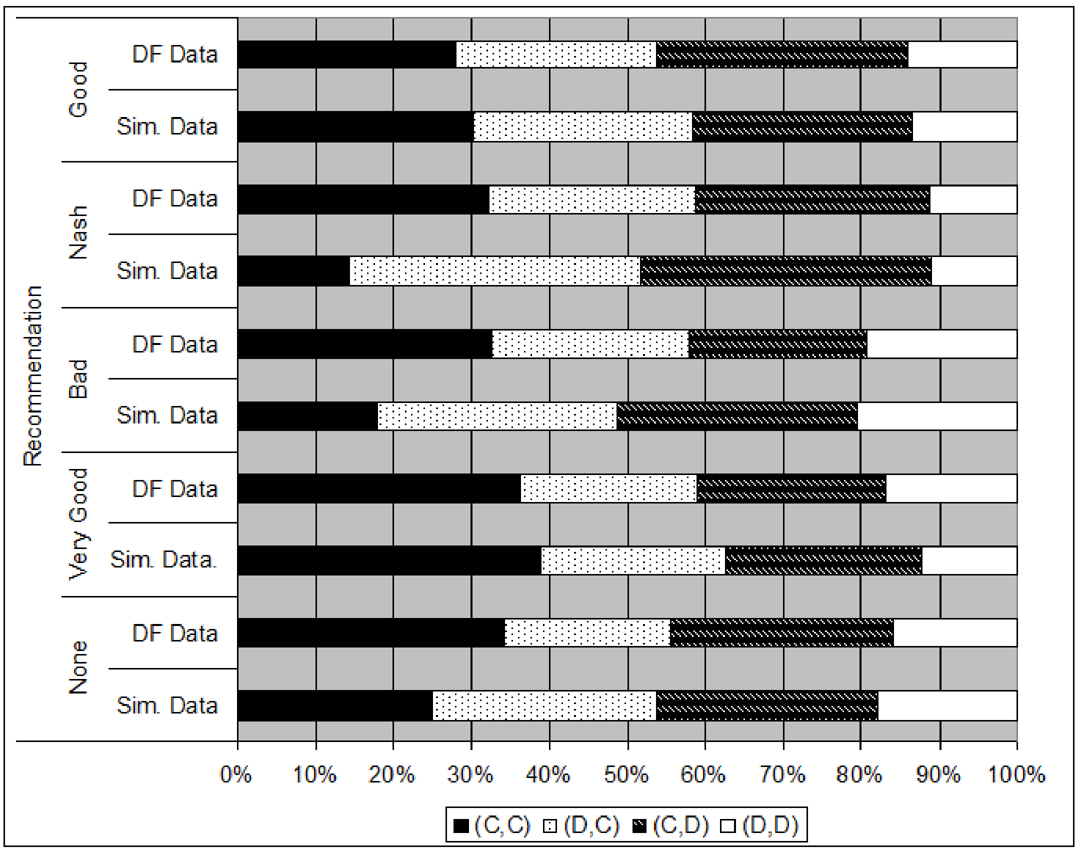


Figure 6: Observed Outcome Frequencies in all Chicken Treatments

Table 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

% 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
	Nash Recommendation								
% Follow	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
	Bad Recommendation								
% Follow	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
	Very Good Recommendation								
% Follow	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 10: Observed Frequencies of Recommendation Following in all Chicken Treatments

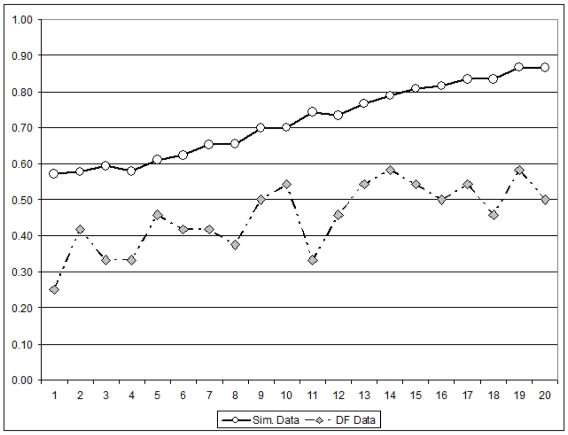
according to whether the recommendation was C or D. The frequency of *pairwise* recommendation following (as studied in the BoS game) is labeled "Pairwise". We report on all of these different measures of recommendation following as we have data on these measures from DF's data set. In Table 10 the mean 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

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

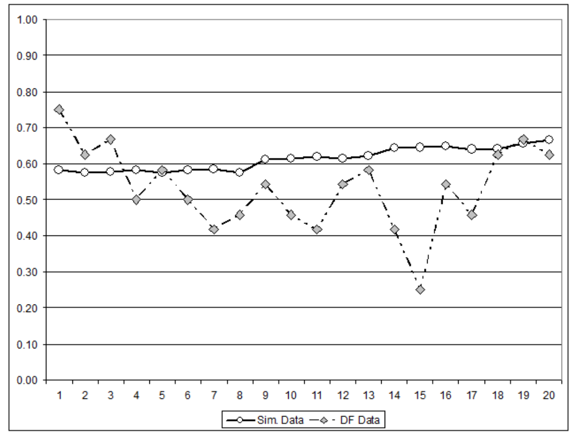
Table 11: IEL nf and f Strategies in all Chicken Treatments

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 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.

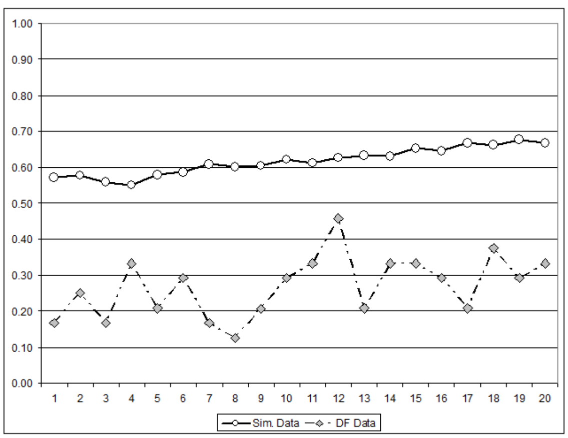
Table 11 reports on the types of strategies the IEL algorithm coordinated on over time, using the strategy classification methodology discussed earlier in connection with Table 4. We see again that the average frequency of recommendation following (the average value of f) is increasing over time in all 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 a



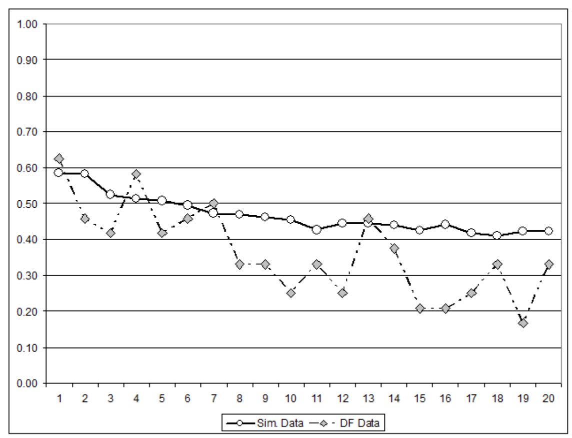
(a) Chicken-Nash



(b) Chicken-Good



(c) Chicken-Bad



(d) Chicken-Very Good

Figure 7: Pairwise Recommendation Following over Time in the Chicken Treatments: Simulations versus DF Data

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 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 10-11, 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.

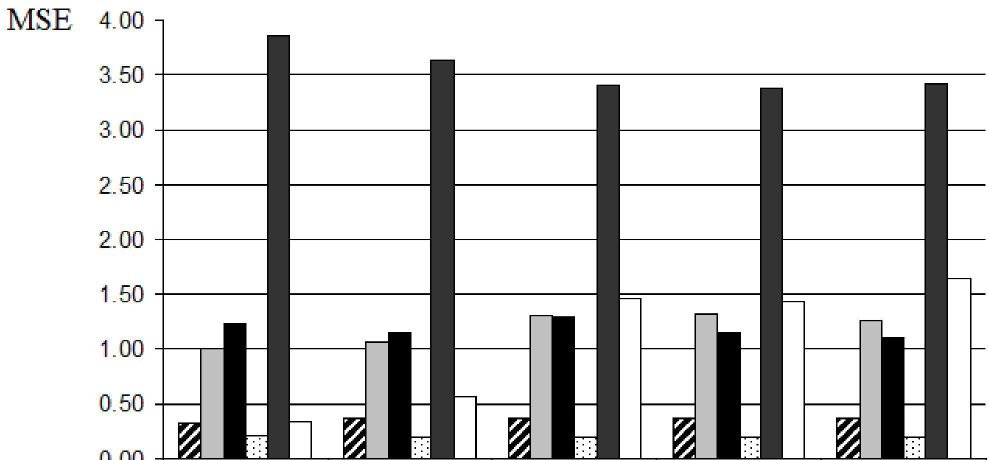
5 Sensitivity Analysis

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 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]

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 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

[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

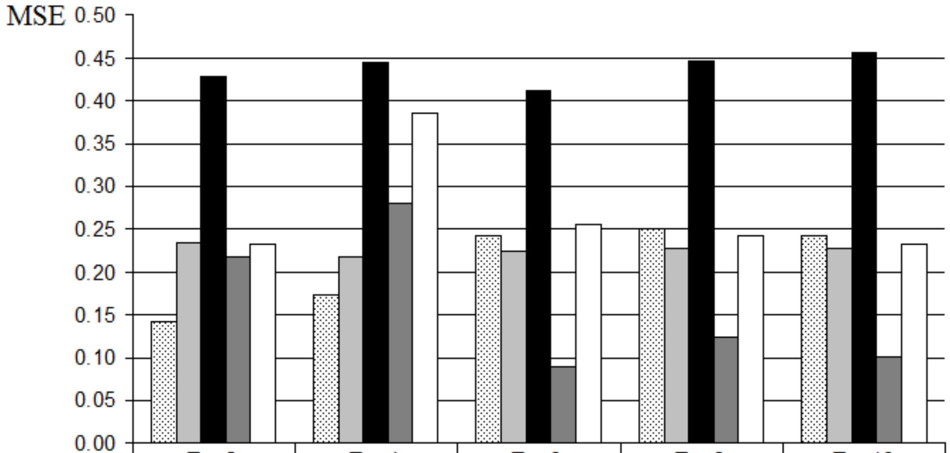
Figure 8: MSE Between Simulations and Experimental Data for Different Evaluation Lengths, Battle of the Sexes Treatments

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 8; analogous results for the Chicken game are shown in Figure 9.

For the Battle of the Sexes game (Figure 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



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

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 $T = 2$ 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 $T = 2$.

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

[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.

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

$$f \sim U\Big[\frac{\psi_{min} - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}{1 - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}, \frac{\psi_{max} - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}{1 - \{\Phi_C(2r_C - 1) + (1 - r_C)\}}\Big] \qquad (3)$$

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.

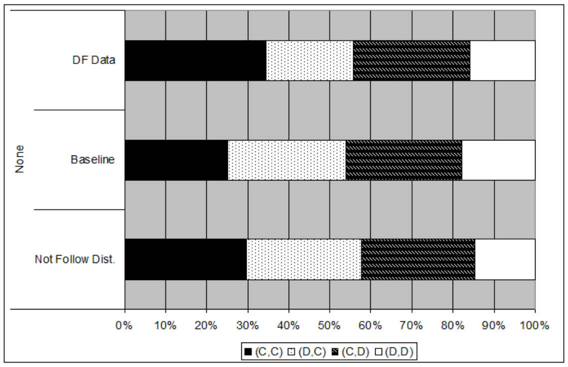
To better understand the marginal contribution of each initialization change, for each Chicken treatment, we conducted 500 runs of the IEL simulation with just the Not Follow part initialized to match the data (i.e., the Follow part remained a uniform random draw over [0,1]) or with just the Follow part 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 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.

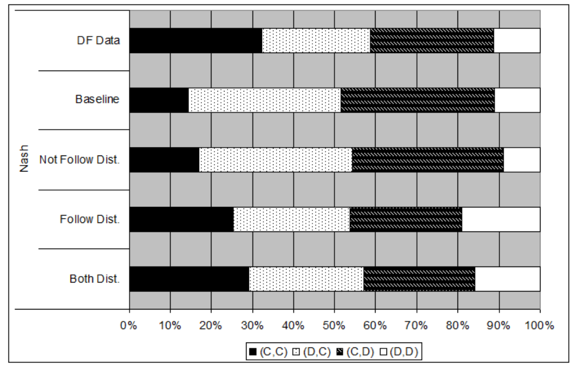
The figures reveal that for some, but not all treatments, a change in initial conditions for the Follow and Not Follow parts of the strategy so as to better match the initial experimental data, can yield improvements in the overall frequencies of outcomes so that these more closely resemble the human subjects data. For instance, in the Chicken-Nash treatment (panel (b) of Figure 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 alway 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.

5.3 Grid Search Description and Results

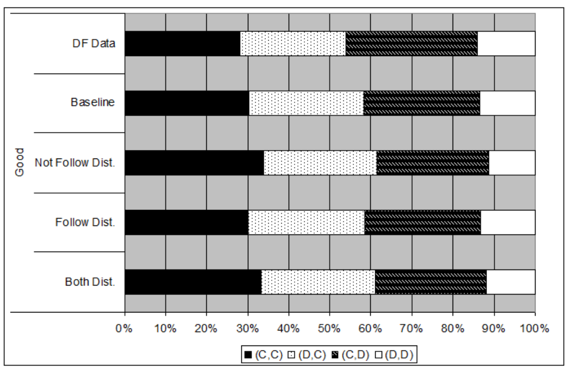
In this section we report on a grid search over the three main free parameters of the IEL with the aim of improving the fit of the model to the experimental data. We again use the mean squared errors of the period-by-period difference in payoffs between the simulated and experimental data to determine the goodness of fit, and we compare our findings with the "Baseline" parameterization used for the IEL simulation results reported in section 4,



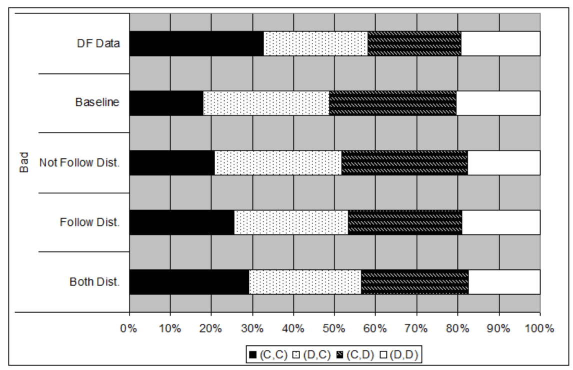
(a) Chicken None



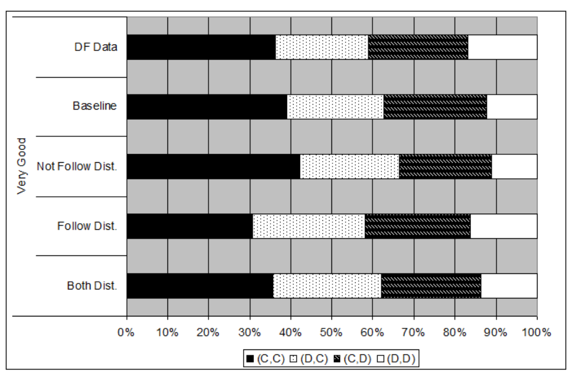
(b) Chicken Nash



(c) Chicken Good



(d) Chicken Bad



(e) Chicken Very Good

Figure 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 A 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.

Tables 12 and 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

	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 12: Comparison MSE Chicken

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 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 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.

[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.

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 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 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. 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 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 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.

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 paper, 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 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.

[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).

References

Arifovic, J. and J. Ledyard (2004). "Scaling Up Learning Models in Public Goods Games." *Journal of Public Economic Theory*, 6, 205-238.

Arifovic, J. and J. Ledyard (2007). "Call Market Book Information and Efficiency." *Journal of Economic Dynamics and Control*, 31, 1971-2000.

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

Arifovic, J. and J. Ledyard (2012). "Individual Evolutionary Learning, Other Regarding Preferences, and the Voluntary Contribution Mechanism." *Journal of Public Economics*, 96, 808-823.

Aumann, R. (1974). "Subjectivity and Correlation in Randomized Strategies." *Journal of Mathematical Economics* 1, 67-96.

Aumann, R. (1987). "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica* 55, 1-18.

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

Boitnott, J.F. (2015). "The Behavior of Learning Algorithms in Experimental Games with Congestion," working paper, Simon Fraser University.

Bone, J., M. Drouvelis and I. Ray (2013). "Coordination in 2 × 2 Games by Following Recommendations from Correlated Equilibria." Working Paper, University of Birmingham.

Cason, T.N., and T. Sharma (2007). "Recommended Play and Correlated Equilibria." *Economic Theory*, 33, 11-27.

Duffy, J. and N. Feltovich (2010). "Correlated Equilibria, Good and Bad: An Experimental Study." *International Economic Review*, 51, 701-721.

Duffy, J., E.K. Lai and W. Lim (2016). "Coordination via Correlation: An Experimental Study." *Economic Theory*, forthcoming.

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

Foster, D.P. and R.V. Vohra (1997). "Calibrated Learning and Correlated Equilibrium." *Games and Economic Behavior*, 21, 40-55.

Fudenberg, D. and D. Levine (1999). "Universal Conditional Consistency." *Games and Economic Behavior*, 29, 104-130.

Gintis, H. (2009). *The Bounds of Reason*. Princeton: Princeton University Press.

Hart, S. and A. Mas-Colell (2000). "A Simple Adaptive Procedure Leading to Correlated Equilibrium", *Econometrica*, 68, 1127 -1150.

Kuzmics, C., T. Palfrey, and Brian W. Rogers (2014). "Symmetric Play in Repeated Allocation Games." *Journal of Economic Theory*, 154, 25-67.

Lehrer, E. (2003). "A Wide Range No-Regret Theorem." *Games and Economic Behavior*, 42, 101-115.

Moreno, D. and J. Wooders (1998). "An Experimental Study of Communication and Coordination in Noncooperative Games," *Games and Economic Behavior* 24, 47-76.

Myerson, R. (1991). *Game Theory: Analysis of Conflict*. Cambridge: Harvard University Press.

Appendix: Results From Simulating the Model with the Parameter Vector Found in the Grid Search (Not Intended for Publication

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 3-4, Figures 1-2, Tables 5-6, Figures 3, Tables 7-8, and Figures 4-5 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 15: 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 16: IEL nf and f Strategies in BoS-DR, Best-Fit Parameter Vector

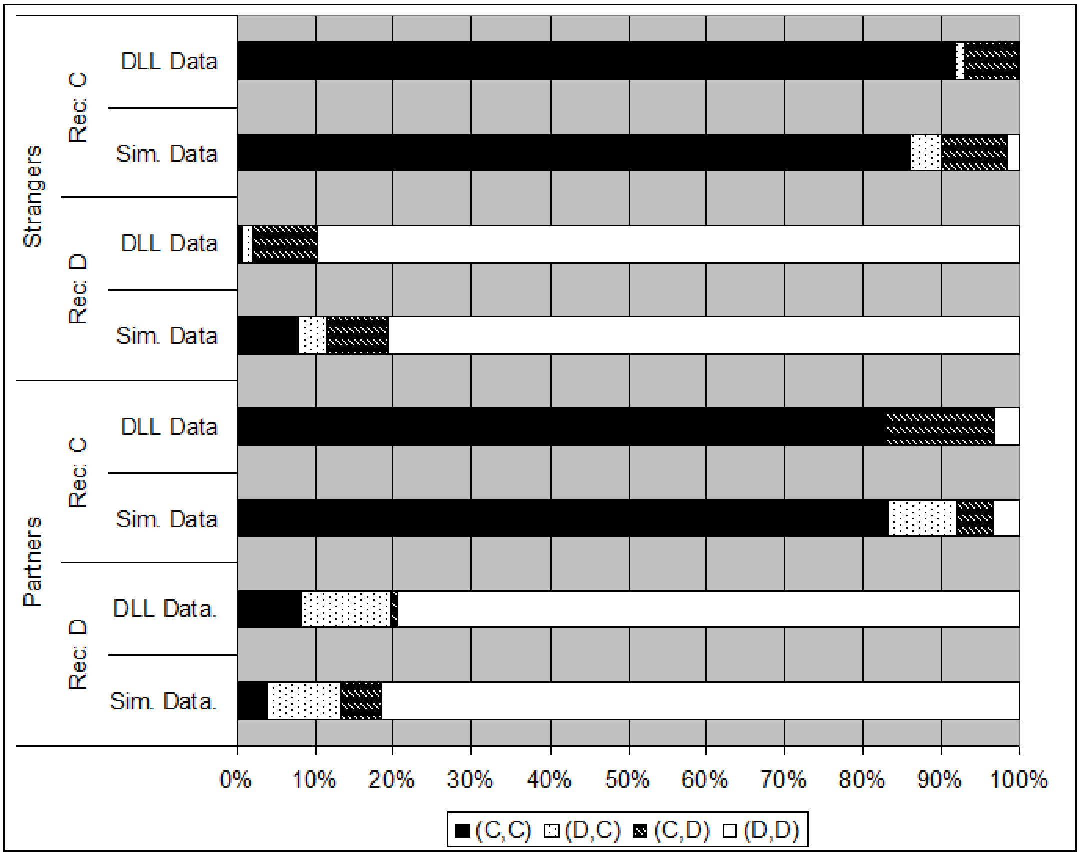
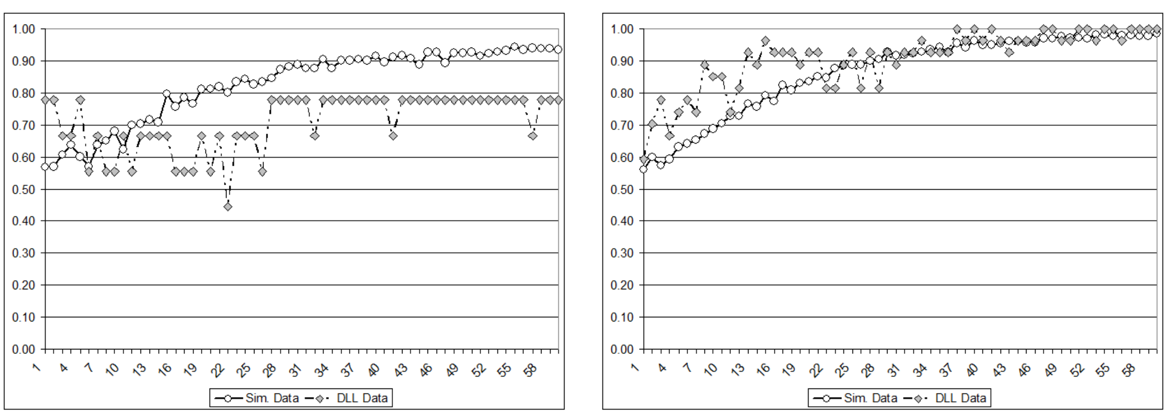


Figure 11: Outcome Frequencies Based on Recommendations in BoS-DR, Best Fit Parameter Vector



(a) BoS-DR Partners

(b) BoS-DR Strangers

Figure 12: 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 17: 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 18: IEL nf and f Strategies in BoS-None, Best-Fit Parameter Vector

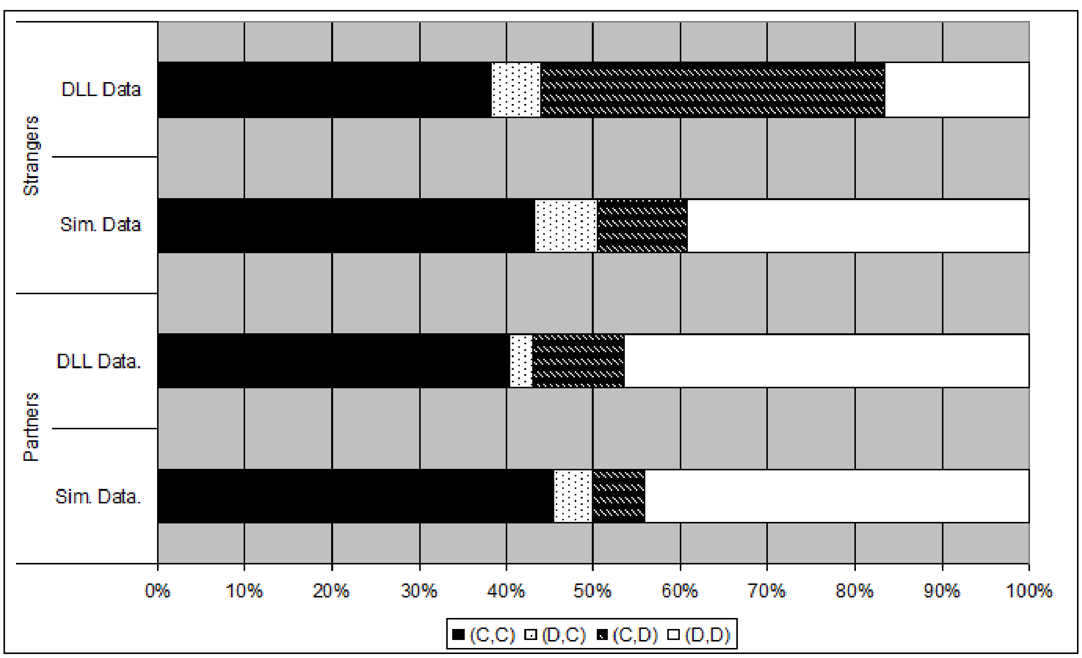


Figure 13: Observed Outcome Frequencies in BoS-None, Best-Fit Parameter Vector

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 19: 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 20: IEL nf and f Strategies in BoS-IR, Best-Fit Parameter Vector

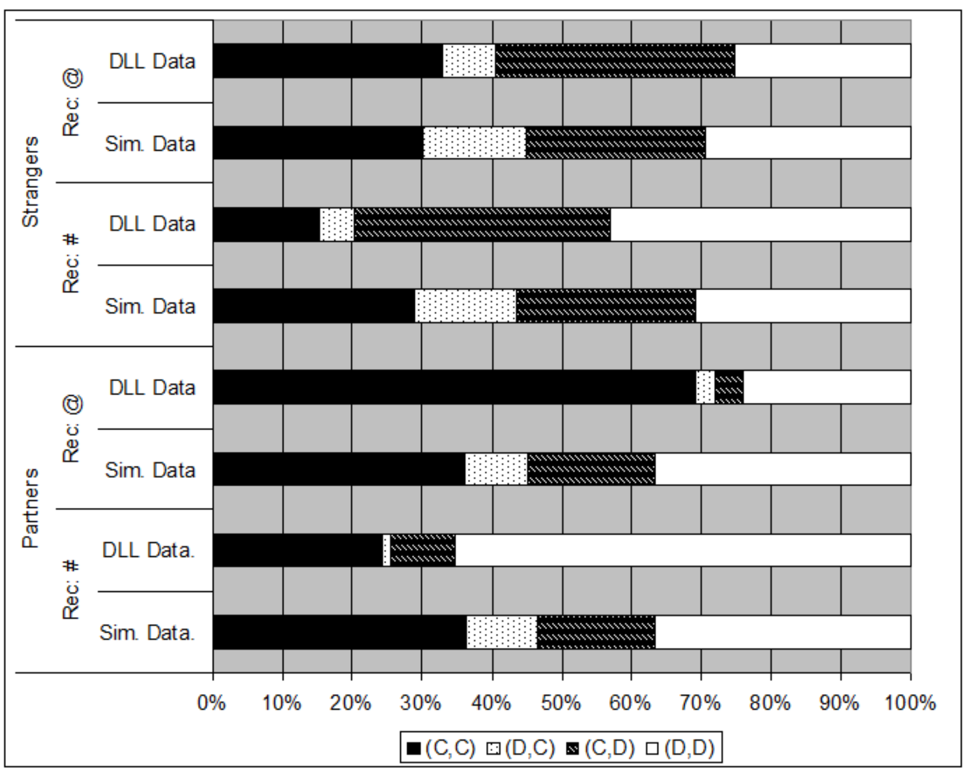
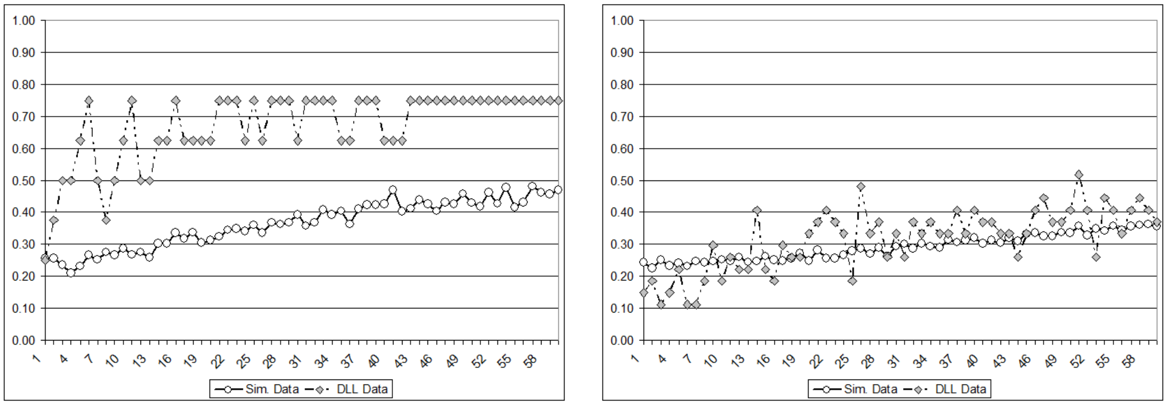


Figure 14: Observed Outcome Frequencies in BoS-IR, Best-Fit Parameter Vector



(a) BoS-IR Partners

(b) BoS-IR Strangers

Figure 15: 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 9-11 and Figures 6-7 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 21: Comparison of Efficiency Measures in all Chicken Treatments, Best-Fit Parameter Vector

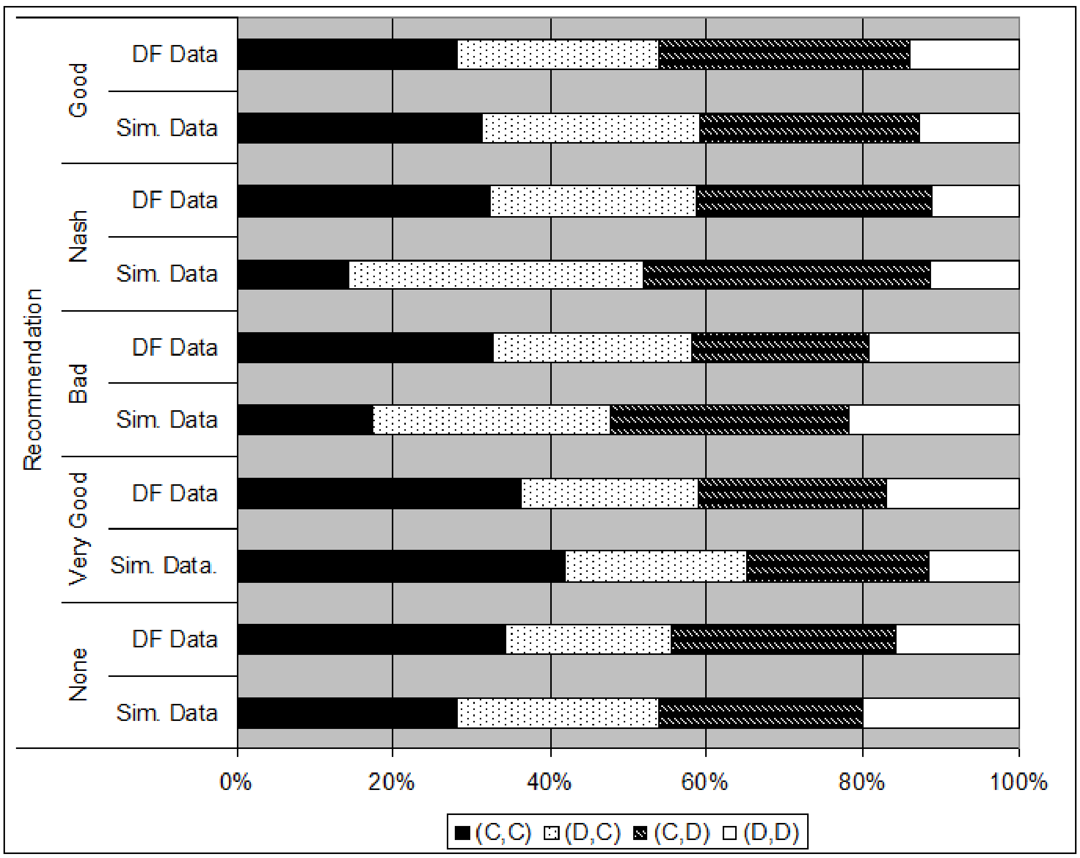


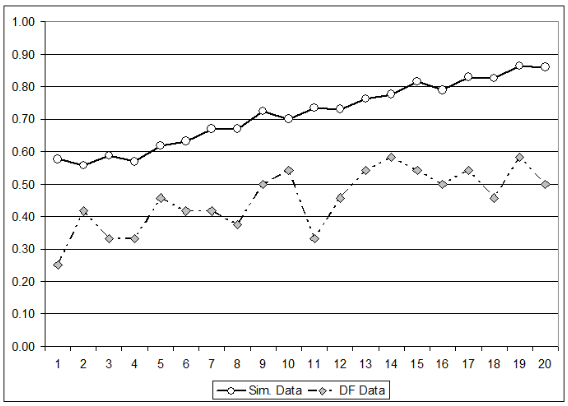
Figure 16: Observed Outcome Frequencies in all Chicken Treatments, Best-Fit Parameter Vector

% 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
	Nash Recommendation								
% Follow	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
	Bad Recommendation								
% Follow	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
	Very Good Recommendation								
% Follow	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

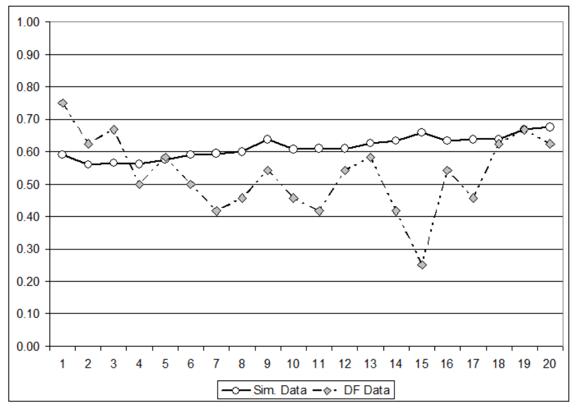
Table 22: Observed Frequencies of Recommendation Following in all Chicken Treatments, Best-Fit Parameter Vector

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

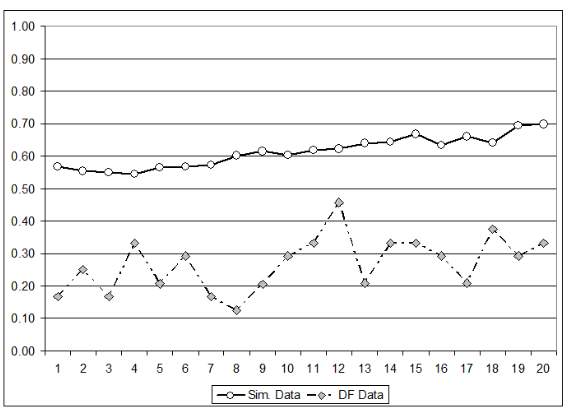
Table 23: IEL nf and f Strategies in all Chicken Treatments, Best-Fit Parameter Vector



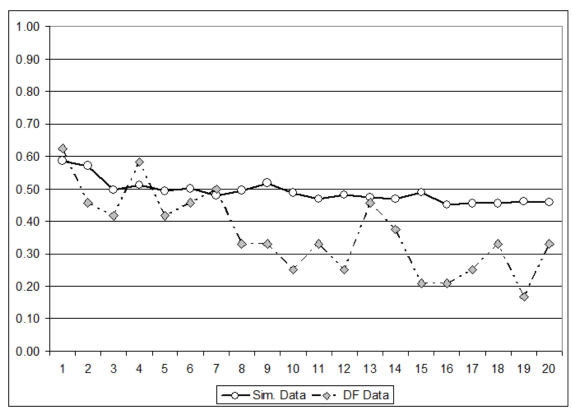
(a) Chicken-Nash



(b) Chicken-Good



(c) Chicken-Bad



(d) Chicken-Very Good

Figure 17: Pairwise Recommendation Following over Time in the Chicken Treatments: Simulations versus DF Data, Best-Fit Parameter Vector