Three Essays on
Bounded Rationality and Individual Learning
in Repeated Games

by

Duncan Whitehead

BSc Economics & Mathematics (Joint Honours) - University of St Andrews, 1996
MSc Economics (SGPE) - University of Edinburgh, 2003

Doctor of Philosophy

in

Economics

in the

College of Humanities and Social Science
of the
UNIVERSITY OF EDINBURGH

Supervisors:
Professor Ed Hopkins (Principal)
Professor József Sákovics (Second)

August 2008

Three Essays on

Bounded Rationality and Individual Learning

in Repeated Games

by

Duncan Whitehead

Contents

Declaration

I certify that this thesis does not incorporate any material previously submitted for a degree or diploma in any University; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person where due reference is not made in the text.

I also certify that the thesis has been composed by myself and that all the work is my own.

27/3/2009

Duncan Whitehead, August 2008

Acknowledgements

I am profoundly indebted to my supervisor, Ed Hopkins, for all he has taught me, for his constant attention, encouragement, and, most of all, patience. I am very grateful to the Edinburgh School of Economics for providing me with a stimulating environment and to all those who provided comments and suggestions. I am also grateful to the Economic & Social Research Council for their financial support.

I would like to thank in particular Jim Engle-Warnick and Julie Héroux for their research assistance with my experiment in economics. I thank the Social Science and Humanities Research Council, the Centre for Interuniversity Research and Analysis on Organizations (CIRANO), and the Royal Economic Society for their financial support of my experiment and the Bell University Laboratory in Electronic Commerce and Experimental Economy for the use of the experimental laboratory.

Last, but not least, I would like to thank my family, for not thinking that I was wasting my time; all my friends, for their faith in me; and especially Kerry and Alex, for tolerating all my anxieties.

Abstract

This thesis is composed of three chapters, which can be read independently.

In the first chapter, we revisit the El Farol bar problem developed by Brian W. Arthur (1994) to investigate how one might best model bounded rationality in economics. We begin by modelling the El Farol bar problem as a market entry game and describing its Nash equilibria. Then, assuming agents are boundedly rational in accordance with a reinforcement learning model, we analyse long-run behaviour in the repeated game. We then state our main result. In a single population of individuals playing the El Farol game, reinforcement learning predicts that the population is eventually subdivided into two distinct groups: those who invariably go to the bar and those who almost never do. In doing so we demonstrate that reinforcement learning predicts sorting in the El Farol bar problem.

The second chapter considers the long-run behaviour of agents learning in finite population games with random matching. In particular we study finite population games composed of anti-coordination pair games. We find the set of conditions for

the payoff matrix of the two-player pair game that ensures the existence of strict pure strategy equilibria in the finite population game. Furthermore, we suggest that if the population is sufficiently large and the two-player pair games meet certain criteria, then the long-run behaviour of individuals, learning in accordance with the Erev and Roth (1998) reinforcement model, asymptotically converges to pure strategy profiles of the population game. These are equilibria where all individual agents play pure strategies, while in aggregate the frequencies of pure strategies played in the population mimic the mixed strategy equilibrium in the pair game. In addition we gather further evidence through computer simulations.

The third chapter investigates some of the theoretical predictions of learning theory in anti-coordination finite population games with random matching through laboratory experiments in economics. Previous data from experiments on anti-coordination games has focused on aggregate behaviour and has evidenced that outcomes mimic the mixed strategy equilibrium. Here we show that in finite population anti-coordination games, reinforcement learning predicts sorting; that is, in the long-run, agents play pure strategy equilibria where subsets of the population permanently play each available action.

List of Figures

Chapter 1

The El Farol Bar Problem Revisited: Reinforcement Learning in a Potential Game

1.1 Introduction

The El Farol bar problem was introduced by Brian W. Arthur (1994) as a framework to investigate how one models bounded rationality in economics. It was inspired by the El Farol bar in Santa Fe, New Mexico, which offered Irish music on Thursday nights. The original problem was constructed as follows:

> "N people decide independently each week whether to go to a bar that offers entertainment on a certain night. For correctness, let us set N at 100. Space is limited, and the evening is enjoyable if things are not too crowded – specifically, if fewer than 60 percent of the possible 100 are present. There is no sure way to tell the numbers coming in advance; therefore a person or an agent goes (deems it worth going) if he expects fewer than 60 to show up or stays home if he expects more than 60 to go."[1]

Arthur's (1994) preliminary results from the field of computational economics show that the number of people attending the bar converges quickly and then hovers around the capacity level of the resource.

Our contribution to the literature on the El Farol bar problem and theory of learning in games is fourfold. First, we apply the Erev and Roth (1998) model of reinforcement learning to the El Farol framework. We believe the Erev and Roth (1998) model of reinforcement learning is the most appropriate individual learning model to apply in this instance, because in general people who stay at home do not know what payoff they would have received if they had gone to the bar. We then

[1] Arthur (1994), pp 409.

prove analytically that long-run behaviour will converge asymptotically to the set of pure strategy Nash equilibria of the El Farol stage game.[2] In other words the number of people attending the bar converges and then hovers around the capacity level of the resource. Furthermore, models of learning, including reinforcement learning and fictitious play, predict sorting in the El Farol bar problem; that is, in a single population of individuals playing the El Farol game, reinforcement learning predicts that the population is eventually subdivided into two distinct groups: those who invariably go to the bar and those who almost never do.[3]

Second, we demonstrate that the El Farol bar problem may be modelled as a market entry game with boundedly rational reinforcement learners. We build upon the work of Duffy and Hopkins (2005), who have proved that in market entry games, where payoffs are decreasing in a continuous manner with respect to the number of other market entrants, the only asymptotically stable Nash equilibria are those corresponding to pure Nash profiles. Our main result also proves asymptotic convergence to those equilibria corresponding to pure Nash profiles in the market entry game. In addition our result also proves that this is the case when payoffs are decreasing in a discontinuous way with respect to the number of other market entrants.

[2]This is in contrast with Franke's (2003) use of numerical simulations of reinforcement learning applied to the El Farol bar problem.

[3]It is worth noting the distinction between learning models and the theory of learning. The latter can be interpreted as a process while the former captures this dynamic. Thus, any implications from learning models carry to the theory on the assumption that processes have been captured in the model.

Third, Sandholm (2001) has proved that, under a broad class of evolutionary dynamics, behaviour convergences to Nash equilibrium from all initial conditions in potential games with continuous player sets. Sandholm's (2001) convergence results assume that individual behaviour adjustments should satisfy what was termed *positive correlation*; meaning any myopic adjustment dynamic that exhibits a positive relationship between growth rates and payoffs in each population. Our result contributes to this literature by proving that, for the evolutionary dynamics associated with Erev and Roth's (1998) model of reinforcement learning, long-run behaviour converges in potential games with finite sets of players.

Finally, there is a contribution to be made to the extensive literature on the El Farol bar problem and its associated problem, the Minority Game in the field of complex systems.[4] Currently, it would appear that the opportunity to apply convergence results from models of individual learning to situations like those represented by the El Farol bar problem has been overlooked.

We will begin by using the tools of game theory to model the El Farol bar problem as a non-cooperative coordination game in which payoffs are determined by negative externalities. We then model the El Farol bar problem as a repeated market-entry game with boundedly rational agents. Analysis of the stage game will show that there are a large number of Nash equilibria. Therefore, equilibria refinement/coordination becomes problematic. In order to refine the equilibria set,

[4]See http://www.unifr.ch/econophysics/minority/ for research on the Minority Game.

we allow players to learn from experience. The analytical tools developed in Duffy and Hopkins (2005), Hopkins and Posch (2005) and Monderer and Shapley (1996) will be employed to study the predicted outcome of play under the Erev and Roth (1998) model of reinforcement learning.

Reinforcement learning assumes that individuals only have access to the attendance figures of the bar for each week that they attend.[5] The long-run behaviour of agents under this adaptive rule will then be considered, and it will be shown that under this learning process, play will converge to the set of pure strategy Nash equilibria with probability one.

The intuition behind our main result is that in the El Farol bar problem reinforcement learners who do not regularly attend are more often than not disappointed when they do choose to do so. Similarly, those who regularly attend always seem to have a good time, and thus are more likely to attend in the future.

A good way to think about this outcome is to imagine that all players in one week play a mixed strategy. It is quite likely that the bar actually turns out to be busy. Therefore, all agents who attended will be reinforced with the lower payoff. This will reduce their propensity to attend in the future. The following week the probability of the bar being overcrowded will be diminished. Those who do attend will most likely receive high payoff reinforcement from attending and their propensity

[5]However, the results presented here within are easily extended to allow for the more generic set-up where all individuals learn attendance figures whether they attend or not. This is often referred to as hypothetical reinforcement or fictitious play learning.

to attend in the future will increase again while that of the players who stayed away will be reduced. Therefore, we have two positive feedback loops. One causes those who attend regularly to do so more often. The other leads those who stay at home to be more likely to do so in the future. We can therefore see that any mixed strategy Nash equilibrium is asymptotically unstable under the dynamics of Erev and Roth (1998) reinforcement learning.

In Section 1.2 we review the El Farol bar problem as introduced by Arthur (1994). We set out his modelling approach to bounded rationality in the El Farol bar problem and summarise the initial results from his computational experiments. We discuss the use of the inductive thinking approach to modelling bounded rationality, both in the El Farol bar problem and its closely related problem, the Minority Game. We then outline our motivation for the application of the individual learning approach to capturing the bounded rationality of decision makers and suggest a reinforcement learning model for the El Farol framework. In Section 1.3 we introduce our model of the El Farol bar problem, define the El Farol stage game and characterise the set of Nash equilibria, set out in detail the Erev and Roth (1998) model of reinforcement learning within the El Farol framework, and write down an expression for player's expected strategy adjustment. In Section 1.4 we state and prove our main result; that in the El Farol bar problem a population of boundedly rational agents who behave in accordance with the Erev and Roth (1998) reinforcement learning model are sorted into those who always attend the El Farol bar and those who always stay at home.

Finally, we provide some concluding remarks in Section 1.5.

1.2 The El Farol Bar Problem

The El Farol bar problem was created by Arthur (1994) as a device to investigate how one might best model bounded rationality in economics. It was inspired by the El Farol bar in Santa Fe, New Mexico, which offered Irish music on Thursday nights. The problem is set out as follows: there is a finite population of people and every Thursday night all of the them want to go to the El Farol bar. However, the El Farol bar is quite small, and it is not enjoyable to go there if it is too crowded. So much so, in fact, that the following rules are in place:

- If less than 60% of the population go to the bar, those who go have a more enjoyable evening at the bar than they would have had had they stayed at home.

- If 60% or more of the population go to the bar, those who go have a worse evening at the bar than they would have had had they stayed at home.

Unfortunately, it is necessary for everyone to decide at the same time whether they will go to the bar or not. They cannot wait and see how many others go on a particular Thursday before deciding to go themselves on that Thursday.

The important characteristic of the El Farol bar problem is that if there was an obvious method that all individuals could use to base their decisions on, then it would be possible to find a deductive solution to the problem. However, no matter what method each individual uses to decide if they will go to the bar or not, if everyone uses the same method it is guaranteed to fail. Therefore, from the point of view of the individual, the problem is ill-defined and no deductive rational solution exists.

Situations like those represented by the El Farol bar problem highlight two specific reasons why perfect deductive reasoning might fail to provide clear solutions to some theoretical problems. The first is simply a question of the cognitive limitations of the mind. Beyond a certain level of complexity, logical capacity fails to cope. The second is that in complex strategic situations individuals cannot always rely on persons they are interacting with to behave under assumptions of perfect rationality. In situations like the El Farol bar problem, individuals are forced into a world where they must choose their strategies based on guesses of their opponents' likely behaviour. Without objective, well-defined, shared assumptions, these types of problems become ill-defined and cannot be solved rationally.

The question that arises is how does one best model bounded rationality in economics when perfect rationality fails? Given the defining characteristic of the El Farol bar problem, namely that finding a deductive rational solution is impossible, it follows that the problem itself could provide a useful framework to explore models of bounded rationality in general.

1.2.1 Inductive Reasoning in the El Farol Framework

Arthur (1994) notes that there is a consensus among psychologists that in situations that are either complicated and/or ill-defined, humans tend to look for patterns in order to develop internal models on which they can base their decisions. These methods are inherently inductive. In the El Farol bar problem, Arthur (1994) follows this line of thought and postulates that individuals decide whether they will go to the bar or not by employing mental models to predict expected future attendance. In other words they create forecasting models. If an individual using a specific forecasting model predicts attendance to be low then, based on that model, that individual would attend and vice-versa if attendance is predicted to be high.

As previously discussed, and deriving from the ill-defined nature of the El Farol bar problem itself, we can conclude that no forecasting model can be employed by all individuals and be accurate at the same time. We can easily demonstrate this fact by assuming that a forecasting model exists that predicts that the attendance in the coming week, given attendance in past weeks, is going to be high. If all individuals use this forecasting model to base their decisions on, then nobody will go to the bar.[6] This then renders the forecast invalid and implies that there exists no single forecasting model that all individuals can use upon which to base their attendance decisions. No deductive solution exists to this problem.

[6]This is reminiscent of Yogi Berra's famous comment, "Oh, that place. It's so crowded nobody goes there anymore."

The Inductive Thinking Approach

Arthur's (1994) approach to modelling bounded rationality in the El Farol bar problem is to assume that each individual has access to a number of forecasting models which they use to make their decisions. Furthermore, they score and rank these models at the end of each week according to their accuracy in order to determine which particular model they should base their decision on.

Formally, Arthur (1994) imagines that each individual utilises a number of forecasting models, denoted s^k, to predict attendance in the coming week. Each model forecasts attendance for the coming week given the history of attendance over the last d weeks, denoted $d(h_{t-1}) \in D$, where D is the set of all possible attendance profiles for the last d weeks and d is an exogenously fixed parameter. Then, following the disclosure of the number of individuals who attended the El Farol bar on the most recent Thursday night, a score is associated with each forecasting model. Specifically, the score, denoted $U_t(s^k)$, is calculated by computing the weighted average of the score of the same model in the previous week and the absolute difference between the forecasting model's last prediction, denoted $s^k(d(h_{t-1}))$, and the most recent realised turnout, denoted y_t. Equation (1.1) formulises this calculation.[7]

$$U_t\left(s^k\right) = \lambda U_{t-1}\left(s^k\right) + (1-\lambda)\left|s^k\left(d\left(h_{t-1}\right) - y_t\right)\right| \tag{1.1}$$

[7]It should be noted that I have taken specific care to outline the El Farol bar problem and Arthur's proposed model of the problem as he originally formulated it. This has been possible due to the work of Zambrano (2004) who re-analysed Arthur's original code.

In each week the forecasting model with the highest score is referred to as the active predictor. On each Thursday individuals undertake the action of either attending the El Farol bar or not in accordance with their active predictor. If an individual's active predictor forecasts the attendance on the coming evening to be high, then that individual will choose not to go to the bar. Conversely, if the active predictor forecasts attendance to be low, then that individual will deem it worthwhile going to the bar and they will anticipate an enjoyable evening of Irish music. Once all individuals have made their decisions, i.e. whether to attend the El Farol bar or not, they are then informed of the actual turnout at the bar. This information is made know publicly to all individuals. Each individual then realises their payoffs, updates the score for all their available forecasting models, and confirms their active predictor for next Thursday's decision.

Agent-Based Computer Simulations

Arthur (1994) investigated this model of the El Farol bar problem through the use of computational experiments. He designed artificial agents and simulated their dynamic interaction over time.

In Arthur's (1994) computer simulations, as in the original formulation of the problem, the size of the population, N, is set to 100 and the enjoyable capacity of the El Farol bar, C, is set to 60. Arthur (1994) then creates a finite set of diverse

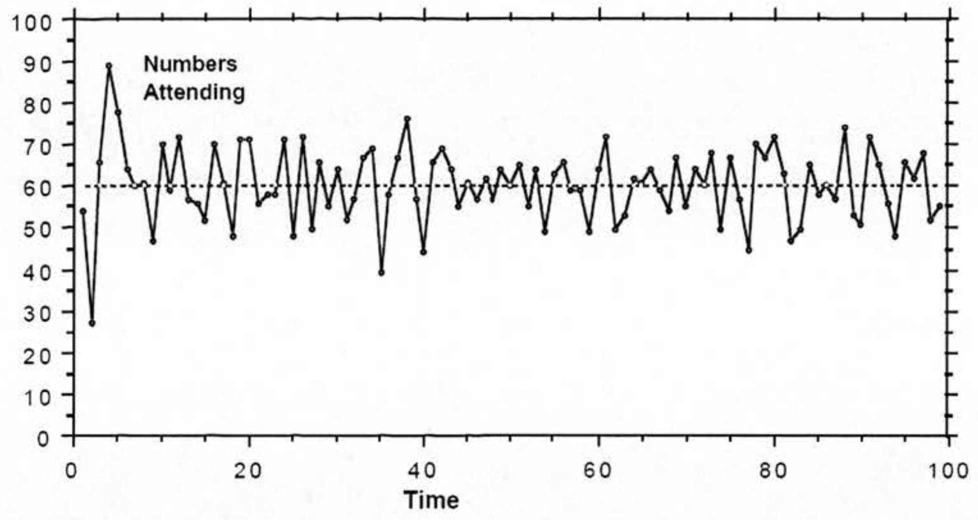

Figure 1.1: Attendance According to Arthur's (1994) Simulations.

forecasting models, or predictors, which map attendance histories to a predicted bar attendance for the coming week. These models were doled out uniformly and randomly, such that each agent was endowed with a non-transferable set of K forecasting models.[8] Each simulation experiment was then run for 100 periods with the combined runs totalling to $10,000$ periods.

The first thing to note about the results of these computer experiments is that, given the starting conditions and the fixed set of predictors available to each simulated agent, the dynamics are completely deterministic. Nevertheless, the simulations produce some interesting results. Two observations become immediately apparent.

[8]This did not preclude the possibility that the agents' predictor sets might overlap.

First, mean attendance always converges to the capacity of the bar. Second, on average 40% of the active predictors forecasted attendance to be higher than the capacity level and 60% below. Arthur (1994) expands on these observations by noting that, "the predictors self organise into an equilibrium pattern or 'ecology'."[9] An example of the attendance rates from a typical run of 100 periods can be seen in Figure 1.1.

The Minority Game

There has been much interest in the El Farol bar problem as a system to study agents in market-like interactions. This has led to the definition of a similar problem called the Minority Game which embodies some basic market mechanisms, while keeping mathematical complexity to a minimum.

The Minority Game is a repeated game where N agents have to decide between two actions, such as buy or sell or attend or not. With N odd this procedure identifies a minority action as that chosen by the minority. Agents who take the minority action are rewarded with one payoff unit. Agents cannot communicate with one another and they have access to publicly available information on the history of past outcomes for a fixed number of periods. As in the El Farol bar problem, the set up requires a prohibitive computational task and, from a strategic point of view, the problem is ill-defined. Again it is postulated that in such complex strategic interactions, agents may

[9]Arthur (1994), pp.409.

prefer to simplify their decision tasks by seeking out behaviour rules, or heuristics, that allocate an action for each possible observed history of outcomes.

The literature on the Minority Game concludes, through both agent-based and analytical models, that there exists a cooperative phase of play when the ratio of the number of unique possible histories to the number of agents, N, is large enough. That is, with respect to the so-called 'random agent' state, in which each agent chooses their action by flipping a coin, agents are better off because the system moves to a sort of 'coordinated' state. The analytical research on the Minority Game employs techniques borrowed from statistical physics in order to describe the game as a spin system, thus enabling the system's properties to be outlined. It should be noted that this avenue of investigation does not enable the study of individual behaviour, but only the system as a whole.

One aspect of this approach, and indeed Arthur's (1994) original investigations, to the El Farol bar problem and bounded rationality is that the theory does not explicitly detail the predictors that should/would be available to each individual/agent. In reality there most likely exists an evolutionary process that regulates the set of predictors as a whole and their availability to each individual agent. Arthur (1994) draws on the following metaphor to make the point: "Just as species, to survive and reproduce, must prove themselves by competing and being adapted within the environment created by other species, in this world hypothesis, to be accurate and therefore acted upon, must prove themselves by competing and being adapted within and environment created by other agents' hypothesis."[10]

1.2.2 Individual Learning in the El Farol Framework

The El Farol bar problem represents a complex strategic environment where rational deductive thinking fails to provide any clear solutions. The question we wish to address is what we should put in place of perfect rationality. In the previous section, we reviewed the literature reporting work that has been directed at achieving this goal within the El Farol framework through the use of inductive reasoning. Suppose instead that individuals in the El Farol bar problem can find their way to an optimal solution by trial and error, i.e. learning.[11] In effect we propose that this is the role that, loosely speaking, the predictors fulfil in Arthur's (1994) original paper on inductive reasoning and bounded rationality in the El Farol bar problem. Recall that if a predictor correctly forecasts attendance, it is more likely to be used as an active predictor. If not, it will not be used. Following this argument it seems reasonable to consider the El Farol bar problem as one with boundedly rational agents who gradually adjust their behaviour over time, until there is no longer any room for improvement in their payoffs.

In game theory the techniques for modelling this type of adaptation process are

[10]Arthur, (1994), pp. 408.

[11]A player cannot adapt to situations that are only encountered once. With this in mind, we must consider players learning equilibria in an identically repeated game environment.

closely related to replicator dynamics. The idea of replicator dynamics was introduced by Maynard Smith (1974) to model dynamic processes in the biological sciences. Essentially, replicator dynamics says that if an individual of a certain type earns an above average payoff, then that individual type's frequency in the population rises. When modelling an individual learning process in a repeated game, we modify this interpretation of replicator dynamics to the following: if an individual who has a propensity to use a particular strategy earns an above average payoff from that strategy, then the propensity to use that strategy in the future increases.

The El Farol bar problem will now be modelled as a repeated market entry game where players adhere to a pre-specified learning process. The manner in which individual learning is modelled in repeated games is simple and quite intuitive. Essentially, individual learning is an algorithm that each player follows in each period of play. Imagine that each individual in the El Farol bar problem, whether they go to the bar or not, keeps an urn by their side. In the urn there are a number of balls coloured either green or red. We can consider these balls to be replacing the function of Arthur's (1994) predictors in the El Farol bar problem.[12] Instead of each individual making their action choice dependent on the forecast of their active predictor, players will choose a ball from their urn and obey its colour coding. In other words if a green ball is selected that individual will go to the bar and if a red ball is chosen they will stay at home. Once a ball is drawn and the corresponding action is taken, the ball is

[12]This is not to be taken literally, but they will provide the same decision function as the predictors do in Arthur's formulation.

then placed back into the urn.

The learning model is then specified by an updating rule. This is the set of instructions that dictates how many balls and of what colours should be added to the urn after each round of play. Using this framework we can describe each player as having propensities for each action. The propensity to undertake a certain action is a function of the number of correspondingly coloured balls in the urn.[13] The probability that a ball of a certain colour will be chosen from a particular individual's urn is determined by the choice rule, which is a mapping from propensities to a number in the unit interval. To find the equilibrium, we calculate in the limit, as the number of repetitions of the game tends to infinity, the probability that each action will be taken.

Let us now recall in detail the motivation for employing an individual learning model of bounded rationality in the El Farol bar problem. As previously stated, the complexity of the problem makes it reasonable to assume that individuals suffer from cognitive limitations. Furthermore, we have already demonstrated that the complexity of beliefs means that, from a strategic viewpoint, individuals are unable to employ deductive reasoning to identify optimal/coordinated strategies. Given these constraints we suppose that individuals find their optimal strategies in the El Farol bar problem through repeated interaction and the application of an adaptive algorithm.

[13] It is also dependent on the choice rule specified in the learning model which shall be expanded on later in the paper.

It will be assumed that any adaptive algorithm will adhere to some basic principles of individual learning.

First, the law of effect: choices that have led to good outcomes in the past are more likely to be repeated in the future. Second, the power law of practice: learning curves should initially be steep and then later they should be flatter. This is paramount to assuming that in any adaptive process the adjustments become smaller over time. Finally, choice behavior should be probabilistic. This is a basic assumption in most mathematical learning theories proposed in psychology. Erev and Roth (1998) have developed a robust model of reinforcement learning which incorporates all these principles that shall be applied to our model of the El Farol bar problem.

1.3 A Model of the El Farol Bar Problem

The El Farol bar problem is essentially a repeated simultaneous move game. There are N players with identical preferences who attempt to coordinate their actions of either going to the bar or staying at home in such a way as to maximise their individual payoffs, subject to the crowding externality from going to the bar. Players need to coordinate their actions, independently and without prior communication, such that:

- when a player decides to go to the bar, i.e. deems it worthy of going to the bar, they can look forward to a payoff that is greater than what they would have received had they stayed at home and
- when a player decides to stay at home, i.e. deems it not worthy of going to the bar, they can look forward to a payoff that is greater than what they would have received had they not stayed at home.

The El Farol bar problem can be interpreted as a market entry game (Franke 2003). In general market entry games are interpreted as truncated two-stage games (Selten and Güth 1982). In the first stage, players simultaneously choose either to enter or stay out of the market. Then, in the second stage, the payoffs of the entrants are determined from their market actions. Usually these payoffs are negatively related to the number of market entrants in a continuous way. However, in the El Farol bar problem, payoffs to players entering the bar are related to the number of bar attendants in a discontinuous manner.

Alternatively, the El Farol bar problem may be viewed as a congestion model and thus can be modelled, a la Rosenthal (1973), as a congestion game.[14] It is a congestion game, because each player's payoff depends on the number of other players who choose to utilise the same resource, namely the El Farol bar. This interpretation has been referred to in many studies of the El Farol bar problem in the literature (e.g. Greenwald, Mishra, and Parikh 1998, Bell and Sethares 1999, Bell and Sethares 2001, Bell, Sethares, and Bucklew 2003, Farago, Greenwald, and Hall 2002, Zambrano 2004), but has rarely been developed.

In this paper we shall initially interpret the El Farol bar problem as a market entry game. Later on in our discussions we shall return to the idea of congestion games, because they have important properties that are useful in understanding the long-run behaviour of boundedly rational agents learning in accordance with a reinforcement model in the El Farol bar problem.

1.3.1 The El Farol Stage Game

Let C, a positive no-zero integer, represent the capacity of the bar. If less than C players choose to go to the bar, then the payoff they receive is allied with the notion

[14]Clearly market entry games are a subset of the larger class of congestion games.

that *ex post* those players deemed it worthwhile going. They receive a payoff strictly greater than the payoff they would have received had they stayed at home. On the other hand, if C or more players choose to go to the bar, then the payoff the bar entrants receive is allied with the notion that, *ex post,* those players did not deem it worthwhile going to the bar. In other words they receive a payoff strictly less than the payoff they would have received had they stayed at home.

		State	
		Uncrowded	Crowded
Player *i*	Go to the Bar	***G***	***B***
	Stay at Home	***S***	***S***

where $G > S > B$

Figure 1.2: State Dependent Payoff for Player i in the El Farol stage game.

The payoff function for each player i consists of an unconditional payoff for staying at home, denoted by S, and a conditional payoff, denoted by G or B, dependent on the state of the bar. There are two states of the bar, crowded or not crowded, and the state is determined by the remaining $N - 1$ players. To ensure the strategic form of the game, the payoffs must be strictly ordered such that $G > S > B$. The payoff structure for representative player i for an isolated Thursday in the El Farol bar problem can be represented by the following payoff matrix (see Figure 1.2).

Given the above preliminaries, we can now define the El Farol stage game as a single-stage market entry game with discontinuous, but weakly monotonic, payoffs in other players' actions.

Definition 1.1 *Define the El Farol stage game as the one shot strategic game* $\Gamma =< N, \Delta, \pi^i >$ *consisting of,*

- N *players indexed by* $i \in \{1, 2, , ..., N\}$,
- *a finite set of actions* $\Delta = \{0, 1\}$ *indexed by* δ, *where* $\delta^i = 1$ *denotes player* i*'s action 'go to the bar' and* $\delta^i = 0$ *denotes player* i*'s action 'stay at home' and*[15]
- *a payoff function* $\pi_i : \delta^i \times \delta^{-i} \to \mathbb{R} = \{S, B, G\}$, *such that* $G > S > B$, *where* $\delta^{-i} = \prod_{i \neq j} \delta^{-j}$ *defines the state of the bar.*

Formally we can write the payoff function as,

$$\pi^i\left(\delta^i\right) = \begin{cases} G & \text{if } \delta^i = 1 \quad \text{and } \sum_{j \neq i} \delta^j < C \\ B & \text{if } \delta^i = 1 \quad \text{and } \sum_{j \neq i} \delta^j \geq C \\ S & \text{if } \delta^i = 0 \end{cases}$$

[15] It should be noted that although we employ the notation Δ to denote the set of only two actions available to each player, we do so only to indicate how the reinforcement learning model would be extended to games with more than two distinct actions.

where $C \in \mathbb{Z}$.

Nash Equilibria in the El Farol Stage Game

Let us now characterise the equilibria of the El Farol stage game. The first thing to note is that the number of Nash equilibria in the El Farol stage game is large and rises quickly as N increases. Furthermore, the number of Nash equilibria is maximised for any given N when $C \approx N/2$. There are essentially three types of Nash equilibria, namely:

- **Pure Strategy Nash Equilibria**

  Nash equilibria where all players play a pure strategy.

- **Symmetric Mixed Strategy Nash Equilibria**

  Nash equilibria where all players play a mixed strategy.

- **Asymmetric Mixed Strategy Nash Equilibria**

  Nash equilibria where some players play a pure strategy and the remaining play a mixed strategy.

Let $\bar{Y}$ denote the set of Nash equilibria of the El Farol stage game. It can be shown that $\bar{Y}$ contains a finite number of elements. In Proposition 1.1 we state the number of pure strategy Nash equilibria, denoted $\bar{Y}_P$. Next, we show via Propositions 1.2 and 1.3 that there exists a unique symmetric mixed strategy Nash equilibrium, denoted $\bar{Y}_S$. And finally in Proposition 1.4, we show that the number of asymmetric mixed strategy Nash equilibria, denoted $\bar{Y}_A$, is countable. Therefore, the number of Nash equilibria in the El Farol stage game is finite.[16]

Proposition 1.1 *The number of pure strategy Nash equilibria in the El Farol stage game with $N \in \mathbb{N}$ players and a capacity of $C \in \mathbb{N}$ is,*

$$\begin{pmatrix} N \\ C \end{pmatrix} = \frac{N!}{C!\,(N-C)!} \tag{1.2}$$

Proof See Section 1.A.1 in Appendix 1.A. ■

The following two propositions together demonstrate that a symmetric mixed strategy Nash equilibrium exists and is unique. In Proposition 1.2 we prove that there is a symmetric mixed strategy Nash equilibrium where all players play the same mixed strategy and that it is unique. In Proposition 1.3 we then prove that if all players are playing a mixed strategy they must be playing the same mixed strategy.

[16]Note that $\bar{Y} = \bar{Y}_P \cup \bar{Y}_S \cup \bar{Y}_A$.

Therefore, we have a unique symmetric mixed strategy Nash equilibrium in the El Farol stage game.[17]

Proposition 1.2 *In the El Farol stage game there is a symmetric mixed strategy equilibrium where all players play the same mixed strategy defined by the strategy tuple* $(\alpha, [1-\alpha])$*, where* α *denotes the probability of going to the bar and* $[1-\alpha]$ *denotes the probability of staying at home. Furthermore,* α *is uniquely defined by the following relationship:*

$$\left(\frac{S-B}{G-B}\right) = \sum_{m=0}^{C-1} \left(\begin{array}{c} N-1 \\ m \end{array} \right) \alpha^m \left[1-\alpha\right]^{N-1-m} \tag{1.3}$$

Proof See Section 1.A.2 in Appendix 1.A. ■

Proposition 1.3 *In a Nash equilibria in the El Farol stage game where all players employ a mixed strategy, all agents must play the same mixed strategy.*

Proof See section 1.A.3 in Appendix 1.A. ■

Let us now consider the asymmetric mixed strategy Nash equilibria. Given that we can calculate the number of pure strategy Nash equilibria from (1.2) and that there is a unique symmetric mixed strategy Nash equilibrium, an approach can be

[17]A similar result has been proved by Cheng (1997).

tabled to demonstrate that the number of asymmetric mixed strategy Nash equilibria is finite.

Proposition 1.4 *The number of asymmetric mixed strategy Nash equilibria in the El Farol stage game is countable.*

Proof See Section 1.A.4 in Appendix 1.A. ■

We have now characterised the Nash equilibria of the El Farol stage game. Furthermore, we have shown that the number of Nash equilibria is finite. This finding will be employed later in proving our main result.

1.3.2 The El Farol Game

For completeness we define the El Farol bar problem as the repeated El Farol stage game with boundedly rational agents who learn in accordance with the Erev and Roth (1998) reinforcement learning model. Let us begin by defining the El Farol game.

Definition 1.2 *The El Farol game is the infinitely repeated El Farol stage game.*

1.3.3 Erev and Roth (1998) Reinforcement Learning

We now set out the procedure for the Erev and Roth (1998) reinforcement learning model in detail.[18] In this learning model, each player i has a propensity to undertake each action in each period, denoted $q_t^i(\delta)$. The timeline of the learning procedure is that in each period t each player i chooses to undertake one of their available actions $\delta \in \Delta = \{0, 1\}$ in accordance with a mapping from the propensities to the unit interval $[0, 1]$. This mapping is defined by the choice rule. The player i then undertakes the action dictated by the choice rule and receives a payoff in that period associated with that action. Player i then updates his propensities. The updating procedure is determined by the updating rule. In the Erev and Roth (1998) reinforcement learning model, the only propensities to be updated are those corresponding to the actual action taken. We can now define the model formally. The learning procedure comprises of three components: the initial conditions, a choice rule and an updating rule.

Initial Conditions

Let $q_t^i(\delta)$ be player i's propensity to play action $\delta \in \Delta$ in period t. In the initial period, $t = 0$, we assume that all players have positive propensities for all possible

[18] It is worth noting that when all payoffs are positive, as they are in our model of the El Farol bar problem, then the Erev and Roth (1998) model of reinforcement learning is equivalent to the Roth and Erev (1995) basis model (without modifications); that is, the Updating Rule/Reinforcement Function are equivalent.

actions. That is,

$$q_t^i(\delta) > 0 \text{ for } t = 0 \text{ and for all } i \in N \text{ and } \delta \in \Delta \tag{1.4}$$

This assumption, along with positive payoffs, will also ensure that $q_t^i(\delta) > 0$ for all t and $\delta \in \Delta$.

Choice Rule

Each player i has positive a propensity, $q_t^i(\delta)$, to take action $\delta \in \Delta = \{0,1\}$ in period t. In models of reinforcement learning, the choice rule provides a mapping from propensities to strategies. Let $(y_t^i, [1 - y_t^i])$ represent player i's mixed strategy in period t with two possible actions $\delta \in \Delta = \{0,1\}$, where y_t^i is the probability placed by agent i on action $\delta = 1$ in period t and $[1 - y_t^i]$ is the probability placed by agent i on action $\delta = 0$ in period t. The choice rule employed in the Erev and Roth (1998) reinforcement learning model is often referred to as the simple choice rule. It is a straightforward probability mapping from propensities to the unit interval $[0,1]$. That is,

$$\Pr(\delta = 1) = y_t^i = \frac{q_t^i(1)}{\sum_{\delta\in\Delta} q_t^i(\delta)} = \frac{q_t^i(1)}{Q_t^i} \tag{1.5}$$

where $Q_t^i = \sum_{\delta\in\Delta} q_t^i(\delta)$.[19]

---

[19]Note that since there are only two possible actions for each player i we can write

$$\Pr(\delta = 0) = \left(1 - y_t^i\right) = \frac{q_t^i(0)}{\sum_{\delta\in\Delta} q_t^i(\delta)} = \frac{q_t^i(0)}{Q_t^i}$$

Updating Rule

Let $\bar{\sigma}^i\left(\delta_t^i, m_t^{-i}\right)$ denote the realised increment to player i's propensity in period t from taking action $\delta \in \Delta = \{0, 1\}$ given the aggregate actions taken by the remaining $N-1$, denoted by m_t^{-i} where $m_t^{-i} = \sum_{j \neq i} \delta_t^j$. To complete, and most crucial to, our reinforcement learning model, we must state the means by which players update their propensities. Specifically, in the Erev and Roth (1998) reinforcement learning model, it takes the form that if agent i takes action δ in period t, then the agent's δth propensity is increased by an increment equal to agent i's realised payoff in that period. All other propensities remain unchanged. In other words only realised payoffs act as reinforcers. We thus have the following updating rule,[20]

$$q_{t+1}^i(\delta) = q_t^i(\delta) + \bar{\sigma}^i\left(\delta_t^i, m_t^{-i}\right) \text{ for all } \delta \in \Delta = \{0, 1\} \tag{1.6}$$

1.3.4 Reinforcement Learning in the El Farol Game

We will now model the El Farol bar problem as the El Farol game with boundedly rational agents who learn according to the Erev and Roth (1998) reinforcement learning model. To study the long-run dynamics of the El Farol game with bounded

[20]Note that this updating rule reveals why in this model of reinforcement learning all payoffs must be positive. Otherwise, there would be a possibility of propensities becoming negative and thus leading to choice probabilities that are undefined.

rational agents learning in accordance with the Erev and Roth (1998) reinforcement model, we need to first write the expected motion of the ith player's $\delta = 1$ strategy adjustment. In order to accomplish this task, we must first define player i's expected payoff increment.

Let $\hat{\sigma}^i\left(\delta_t^i, y_t^{-i}\right)$ denote the expected increment to player i's propensity in period t from taking action δ given the aggregate actions taken by the remaining $N-1$ players, denoted by y_t^{-i}, where y_t^{-i} is a vector strategy profile. Note that the updating rule in the Erev and Roth (1998) reinforcement learning model is a function of realised payoffs. However, the expected motion of the ith player's $\delta = 1$ strategy adjustment will be a function of expected payoff increments. This is quantitatively and qualitatively different from realised payoff increments.

Expected Strategy Adjustment in the El Farol Game

To obtain analytical results from the application of Erev and Roth (1998) reinforcement learning model to the El Farol game, we make use of results from the theory of stochastic approximation. In essence we investigate the behaviour of the stochastic learning model by evaluating its expected motion as $t \longrightarrow \infty$. In the case of the Erev and Roth (1998) learning model defined by the choice rule (1.5) and updating rule (1.6), we can write down the expected motion of the ith player's $\delta = 1$ strategy adjustment through the following proposition:

Proposition 1.5 *Given the choice rule (1.5) and the updating rule (1.6), the expected motion of the ith player's $\delta = 1$ strategy adjustment in the repeated El Farol game is:*

$$E\left[y_{t+1}^i | y_t^i\right] - y_t^i = \frac{1}{Q_t^i} y_t^i \left[1 - y_t^i\right] \left[\hat{\sigma}^i\left(1, y_t^{-i}\right) - \hat{\sigma}^i\left(0, y_t^{-i}\right)\right] + O\left(\frac{1}{\left[Q_t^i\right]^2}\right) \tag{1.7}$$

Proof See Section 1.B.1 in Appendix 1.B. ■

1.4 Long-run Behaviour in the El Farol Game

We now arrive at our main result. We consider the behaviour of the expected motion of the players' $\delta = 1$ strategy adjustment as $t \longrightarrow \infty$. We begin by stating the main result.

Theorem 1.1 (Main Result) *If agents in the repeated El Farol game as defined employ the choice rule (1.5) and reinforcement updating rule (1.6) for all of $N \in \mathbb{N}$ and $C \in \mathbb{N}$ such that $C \leq N - 1$ and payoffs such that $G > S > B \geq 0$, with probability one the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot El Farol game. That is,*

$$\Pr \left\{ \lim_{t \longrightarrow \infty} y_t \in \bar{Y}_P \right\} = 1,$$

where $y_t = \left\{ y_t^1, y_t^2, ..., y_t^N \right\}$, $y_t \in Y$, is a strategy profile for the N agents and $\bar{Y}_P$ is the set of pure Nash equilibrium profiles.

Now we prove the main result, Theorem 1.1. In the El Farol game with identical boundedly rational agents, learning according to the Erev and Roth (1998) reinforcement learning model, long-run behaviour converges asymptotically to the set of pure strategy Nash equilibria of the El Farol stage game. This result is established by studying the convergent behaviour of the discrete time stochastic process (1.7) describing the expected strategy adjustment of player i's action of going to the bar. In essence we wish to investigate the limit of this process as $t \to \infty$.

We accomplish this task in two main stages: a positive convergence statement and a negative one. Drawing these two results together we prove our main result. Each stage employs results from the literature on stochastic approximation. First, a result of Benaïm (1999, Corollary 6.6) is employed to demonstrate that the stochastic process will, in the limit as $t \to \infty$, converge asymptotically to one of the fixed points of the adjusted replicator dynamics. Second, two results of Hopkins and Posch (2005, Proposition 2 and 3) are utilised to demonstrate that the stochastic process describing the expected strategy adjustment of player i's action of going to the bar will *not* converge asymptotically to any fixed points that do *not* correspond to a Nash equilibria of the El Farol stage game or to any corresponding Nash equilibria that are unstable under the adjusted replicator dynamics. These two stages combined will imply that the discrete time stochastic process describing the expected strategy adjustment of player i's action of going to the bar converges asymptotically to the set of pure strategy equilibria of the El Farol stage game.

1.4.1 Proof of Main Result: First Stage

In the first stage of the proof, we show that the discrete time stochastic process (1.7) converges with probability one to one of the fixed points of the standard replicator dynamics.

Consider for a moment the behaviour of the following stochastic process (Benveniste, Métivier, and Priouret 1990):

$$x_{t+1} - x_t = \gamma_t f(x_t) + \gamma_t \eta_t(x_t) + O\left([\gamma_t]^2\right) \tag{1.8}$$

where x_t lies in $\mathbb{R}^N$, $E[\eta_t(x_t)\,|x_t] = 0$ and γ_t defines the nature of the gain in this adaptive process. For our purposes γ_t is interpreted as the step size of the learning algorithm. In our analysis we wish to study the generic convergence properties of stochastic processes of this form as $t \longrightarrow \infty$.

It turns out that the nature of the gain is important in determining what inferences can be made about the behaviour of (1.8) in the limit. In fact the stronger results from the theory of stochastic approximation apply to adaptive algorithms with decreasing gain, that is stochastic processes with decreasing step size.

Definition 1.3 *The stochastic process (1.8) is said to have decreasing gain if*

$$\sum_t (\gamma_t)^\alpha < \infty \textit{ for some } \alpha > 1 \textit{ where } \sum_t \gamma_t = +\infty$$

For example a common step size of $\gamma_t = 1/t$ would ensure that (1.8) has decreasing gain. It emerges that as $t \longrightarrow \infty$ there is a close relationship between the behaviour of stochastic processes (1.8) with a decreasing gain and the mean or averaged ordinary differential equation of the stochastic process.

$$\dot{x} = f(x) \tag{1.9}$$

In particular it can be shown via Benaïm (1999, Corollary 6.6) that if (1.9) meets certain criteria, the stochastic process (1.8) must converge with probability one to one of the fixed points of the mean or averaged ordinary differential equation (1.9).

Theorem 1.2 (Benaïm (1999, Corollary 6.6)) *If the dynamic process (1.9) admits a strict Lyapunov function and processes a finite number of fixed points, then with probability one the stochastic process (1.8) converges to one of these fixed points.*

We now have a method of illustrating that the long-run behaviour of boundedly rational agents, adjusting their strategies according to the Erev and Roth (1998) reinforcement learning model, in the El Farol game converges to one of the fixed points of mean or averaged differential equation (1.9) associated with the vector of player's expected strategy adjustments.

In order to apply this general result, we must first identify the mean or averaged differential system associated with players' expected strategy adjustment. Furthermore, it must be shown that the mean or averaged differential system admits a strict Lyapunov function. And finally, we must establish that the mean or averaged differential system possesses a finite number of isolated fixed points. In the next three subsections we purport to demonstrate just that.

The Joint Dynamic System

One might hope that the standard replicator dynamics represent the mean or averaged differential system derived from the discrete time stochastic process (1.7).

$$\dot{y}^i = y^i \left[1 - y^i\right] \left[\hat{\sigma}^i \left(1, y^{-i}\right) - \hat{\sigma}^i \left(0, y^{-i}\right)\right] \tag{1.10}$$

Unfortunately, the standard replicator dynamics (1.10) do not for two simple reasons. First, in the Erev and Roth (1998) model, the step size is endogenous; that is, it is determined by the accumulation of payoffs, and thus is not exogenously fixed. Second, the step size is not a scalar.

In order to account for these discrepancies, let us introduce a common step size of $\gamma_t = 1/t$ and N new variables μ_t^i, such that:

$$\mu_t^i = \frac{t}{Q_t^i}$$

We can now substitute $\gamma_t \mu_t^i$ for $1/Q_t^i$ in our discrete time stochastic process (1.7) and arrive at the following *corrected* expected motion of the ith player's strategy adjustment of going to the bar:

$$\begin{aligned} E\left[y_{t+1}^i | y_t^i\right] - y_t^i = {} & \gamma_t \mu_t^i y_t^i \left[1 - y_t^i\right] \left[\hat{\sigma}^i \left(1, y_t^{-i}\right) - \hat{\sigma}^i \left(0, y_t^{-i}\right)\right] \\ & + \gamma_t \xi_t \left(y_t^i\right) + O\left(\left[\gamma_t\right]^2\right) \end{aligned} \tag{1.11}$$

Since we have assumed that all payoffs in the El Farol game are positive to ensure that choice probabilities are well defined, it follows that μ_t^i is bounded away from zero. Furthermore, since $\mu_t^i = t/Q_t^i$ equals the inverse of the average payoff in the limit as $t \to \infty$, it follows that the associated mean or averaged differential equation (1.9) associated with the corrected discrete time stochastic process (1.11) is:

$$\dot{y}^i = \mu^i y^i \left[1 - y^i\right] \left[\hat{\sigma}^i \left(1, y^{-i}\right) - \hat{\sigma}^i \left(0, y^{-i}\right)\right] \tag{1.12}$$

In equilibrium this amounts to the standard definition of the adjusted replicator dynamics. This is extremely useful because there are many results in the literature on the equilibrium behaviour of the adjusted replicator dynamics (see Fudenberg and Levine 1998, Hopkins 2002). We shall revisit some of these findings later in this proof of the main result.

Because each μ_t^i varies over time, we require a further set of N equations describing the discrete time stochastic process of μ_t^i. Using the method we previously employed to write player i's expected strategy adjustment of going to the bar, we now find the expected change player i's step size.

Lemma 1.1 *Given the choice rule (1.5) and the updating rule (1.6), the expected motion of the ith player's step size in the El Farol game is:*

$$\begin{aligned} E\left[\mu_{t+1}^i | \mu_t^i\right] - \mu_t^i = \gamma_t \mu_t^i - \gamma_t \left[\mu_t^i\right]^2 \hat{\sigma}^i \left(0, y_t^{-i}\right) \\ + \gamma_t \left[\mu_t^i\right]^2 y_t^i \left[\hat{\sigma}^i \left(0, y_t^{-i}\right) - \hat{\sigma}^i \left(1, y_t^{-i}\right)\right] \\ + \gamma_t \xi_t \left(y_t^i\right) + O\left(\left[\gamma_t\right]^2\right) \end{aligned} \tag{1.13}$$

Proof First, imagine that player i chooses to attend the bar in period t. The expected change in the player step size can be written as:

$$\mu_{t+1}^i - \mu_t^i(t) = \frac{t+1}{Q_t^i + \hat{\sigma}^i(1, y_t^{-i})} - \frac{t}{Q_t^i}$$
$$= 1 - \mu_t^i \hat{\sigma}^i(1, y_t^{-i}) + O(\gamma_t)$$

Now consider the expected change in step size if player i stays at home.

$$\mu_{t+1}^i - \mu_t^i(t) = \frac{t+1}{Q_t^i + \hat{\sigma}^i(0, y_t^{-i})} - \frac{t}{Q_t^i}$$
$$= 1 - \mu_t^i \hat{\sigma}^i(0, y_t^{-i}) + O(\gamma_t)$$

The expected motion of each player i's step size given μ_t^i can now be written as the expected motion in the step size given y_t^i times the step size in period t.

$$E\left[\mu_{t+1}^i | \mu_t^i(t)\right] - \mu_t^i(t) = \gamma_t \mu_t^i y_t^i \left[1 - \mu_t^i \hat{\sigma}^i(1, y_t^{-i}) + O(\gamma_t)\right]$$
$$+ \gamma_t \mu_t^i \left[1 - y_t^i\right] \left[1 - \mu_t^i \hat{\sigma}^i(0, y_t^{-i}) + O(\gamma_t)\right]$$

and after some more algebraic manipulation we arrive at (1.13). ■

The mean or averaged differential equation derived from the discrete time stochastic process (1.7) has now been corrected for the endogenous and non-scalar step size. Therefore, we have the following mean or averaged differential system consisting of $2N$ differential equations with $2N$ endogenous variables:

$$\dot{y}^i = \mu^i y^i \left[1 - y^i\right] \left[\hat{\sigma}^i(1, y^{-i}) - \hat{\sigma}^i(0, y^{-i})\right] \tag{1.14a}$$

$$\dot{\mu}^i = \mu^i \left[1 - \mu^i \left[\hat{\sigma}^i(0, y^{-i}) + y^i \left[\hat{\sigma}^i(0, y^{-i}) - \hat{\sigma}^i(1, y^{-i})\right]\right]\right] \tag{1.14b}$$

Let us refer to this as the joint dynamic system describing the evolution of player i's strategy adjustment of going to the bar in the El Farol game.

Admission of a Strict Lyapunov Function

We must show that the associated mean or averaged ordinary differential system, the joint dynamic system (1.14), admits a strict Lyapunov function. Let us begin with some definitions.

Definition 1.4 *Let (1.9) be an ordinary differential equation defined on some subset Y of $\mathbb{R}^N$. Let $V : Y \to \mathbb{R}$ be a continuously differentiable function. Furthermore, let $\bar{y}$ be a fixed point of $V(y)$. $V(y)$ is a Lyapunov function if,*

$$\dot{V}(y) \geq 0, \qquad \forall \quad y \in Y \text{ and} \tag{1.15a}$$

$$\dot{V}(\bar{y}) = 0, \qquad \forall \quad \bar{y} \in \theta \tag{1.15b}$$

where θ is the set of fixed points of (1.9).

Definition 1.5 *A strict Lyapunov function is a Lyapunov function $V(y)$ such that:*

$$\dot{V}(y) > 0, \qquad \forall \quad y \notin \theta \tag{1.16}$$

In general it can be difficult and time consuming to identify a suitable Lyapunov function for a particular system. It is often a process of trial and error. An approach to this aspect of the problem developed in the existing literature on the convergence of learning models in games (see Duffy and Hopkins 2005) has been to explicitly derive a suitable function for $V(y)$ and then show that it admits a strict Lyapunov function. In theory, but not always in practice, this can be accomplished by first assuming that $V(y)$ indeed admits a strict Lyapunov function. If this is the case, then the partial derivative $\partial V(y)/\partial y^i$ represents the expected payoff increment to player i from going to the bar.

$$\frac{\partial V(y)}{\partial y^i} = \hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right) \tag{1.17}$$

It should then just be a question of integrating $\partial V(y)/\partial y^i$ with respect to y^i in order to find a suitable $V(y)$ and checking that both conditions (1.15) and (1.16) defining strict Lyapunov functions are met.

The difficulty with this approach is in explicitly finding a function $V(y)$. Expressing $\hat{\sigma}^i(1, y^{-i}) - \hat{\sigma}^i(0, y^{-i})$ in a compact form is not as straightforward as one might first hope. This can be demonstrated by examining $\partial V(y)/\partial y^i$ further. Note that (1.17) can be expressed as:

$$\begin{aligned}\frac{\partial V(y)}{\partial y^i} &= E\left[\pi^i|\delta = 1\right] - E\left[\pi^i|\delta = 0\right] \\ &= [B - S] + \phi\,[G - B]\end{aligned}$$

where

$$\phi = \sum_{j=0}^{C-1} \Pr\left(m_t^{-i} = j\right) \tag{1.18}$$

ϕ is the probability that $C-1$ players or less of the remaining $N-1$ players choose to go to the El Farol bar. It is writing out this latter probability expression (1.18) for ϕ that is unfortunately problematic and can get cumbersome very quickly. Therefore, this turns out to be an intractable method of demonstrating that the joint dynamic system (1.14) admits a strict Lyapunov function.

An alternative approach is to employ a result of Monderer and Shapley (1996, Theorem 3.1) from the theory of potential games to demonstrate that the joint system (1.14) admits a strict Lyapunov function. The argument is as follows: the El Farol game is a congestion game therefore it is a potential game and thus admits a potential function. The properties of potential functions are similar to those of strict Lyapunov functions and therefore, it follows that the joint dynamic system (1.14) admits a strict Lyapunov function.

Let us now begin with some definitions and a restating of Monderer and Shapley (1996, Theorem 3.1).

Definition 1.6 *Let $\Gamma\left(N, Y^i, \pi^i\right)$ be a game in strategic form. Γ is called a potential game if it admits a potential function.*

Definition 1.7 *A function $P : Y \to \mathbb{R}$ is a potential function for Γ, if for every $i \in N$ and for every $y^{-i} \in Y^{-i}$*

$$\pi^i\left(x, y^{-i}\right) - \pi^i\left(x', y^{-i}\right) = P\left(x, y^{-i}\right) - P\left(x', y^{-i}\right) \qquad \forall \quad x, x' \in Y^i$$

Theorem 1.3 (Monderer and Shapley (1996, Theorem 3.1)) *Every congestion game is a potential game.*

Now we can show that the joint dynamic system (1.14) admits a strict Lyapunov function.

Lemma 1.2 *The joint dynamic system (1.14) admits a strict Lyapunov function.*

Proof The El Farol stage game is a congestion game and therefore by Theorem 1.3, Monderer and Shapley (1996, Theorem 3.1), it is a potential game. Thus, there exists a function $P : \delta^i \times \delta^{-i} \to \mathbb{R}$ for every $i \in N$ and for every $\delta^{-i} \in \Delta^{-i}$ such that:

$$\pi^i\left(1, \delta^{-i}\right) - \pi^i\left(0, \delta^{-i}\right) = P\left(1, \delta^{-i}\right) - P\left(0, \delta^{-i}\right) \qquad \forall \quad \delta \in \Delta = \{0, 1\}$$

Given that there is a continuous set of mixed strategies, we can write the potential function $P(y)$ as a smooth function with respect to the strategy space $y \in [0,1]^N$. $P(y)$ is therefore continuously differentiable. Therefore, for every $i \in N$ and for every $x^{-i} \in [0,1]^{N-1}$,

$$\pi^i\left(x, y^{-i}\right) - \pi^i\left(x', y^{-i}\right) = P\left(x, y^{-i}\right) - P\left(x', y^{-i}\right) \qquad \forall \quad x, x' \in [0,1]$$

Now choose x and x' equal to 0 and 1 respectively and take expectations of both sides. It follows that for every $i \in N$ and for every $y^{-i} \in [0,1]^{N-1}$,

$$\hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right) = P\left(1, y^{-i}\right) - P\left(0, y^{-i}\right)$$

Or otherwise stated,

$$\frac{\partial P(y)}{\partial y^i} = \hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right) \tag{1.19}$$

Furthermore,

$$\begin{aligned}
\dot{P}(y) &= \frac{dP(y)}{dy^i}\dot{y}^i \\
&= \hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right)\dot{y}^i \\
&= \mu^i y^i\left[1 - y^i\right]\left[\hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right)\right]^2 \geq 0
\end{aligned}$$

By assumption, $\mu^i > 0$ and $y^i \in [0,1]$. Since $\left[\hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right)\right]^2 \geq 0$ we have that $\dot{P}(y)$ is non negative. Additionally, at any fixed point $\bar{y} \in \theta$ either $\bar{y}^i = 0$, $\left(1 - \bar{y}^i\right) = 0$ or $\left[\hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right)\right] = 0$. Thus $P(y)$ admits a Lyapunov function.

At any $y \notin \theta$, $\dot{y}^i \neq 0$. It should be obvious that:

$$\dot{P}(\bar{y}) = \mu^i y^i\left(1 - y^i\right)\left[\hat{\sigma}^i\left(1, y^{-i}\right) - \hat{\sigma}^i\left(0, y^{-i}\right)\right]^2 > 0.$$

Therefore, $P(y)$ admits a strict Lyapunov function. It follows that the joint dynamic system (1.14) admits a strict Lyapunov function. ■

Fixed Points of the Joint Dynamic System

Definition 1.8 *The fixed points of the joint dynamic system (1.14) are defined as* $\bar{x} = (\bar{y}, \bar{\mu})$ *such that* $\dot{y} = 0$ *and* $\dot{\mu} = 0$.

Lemma 1.3 *The joint dynamic system (1.14) possesses a finite number of isolated fixed points.*

Proof Consider the joint dynamic system (1.14). The fixed points of the N equations describing the evolution of the step size occur when either:

$$\bar{\mu}^i = 0, \frac{1}{\left[\hat{\sigma}^i\left(0, y_t^{-i}\right) + y^i\left[\hat{\sigma}^i\left(0, y_t^{-i}\right) - \hat{\sigma}^i\left(1, y_t^{-i}\right)\right]\right]}$$

By assumption, all payoffs are positive therefore $\bar{\mu}^i$ is bounded away from zero. This means that the fixed points of the joint dynamic system (1.14) with $\bar{\mu}^i = 0$ are always unstable (see Hopkins 2002, Duffy and Hopkins 2005) and therefore are never asymptotic outcomes. We can thus concentrate on the latter case.

Consider the first N equations of the joint dynamic system (1.14). Once we substitute for $\bar{\mu}^i$ and multiply both sides by the denominator we have:

$$y^i\left[1 - y^i\right]\left[\hat{\sigma}^i\left(1, y_t^{-i}\right) - \hat{\sigma}^i\left(0, y_t^{-i}\right)\right] = 0$$

In other words the fixed points of the joint dynamic system (1.14) are exactly the same as those under the adjusted replicator dynamics (1.12) and, consequently, the standard replicator dynamics (1.10). The characterisation of the fixed point of the standard replicator dynamics (1.10) is well known (see Weibull 1995) and consists of the union of all pure states and Nash equilibria of the underlying game.

The number of pure states is obviously finite and, as proved in Propositions 1.2-1.4, the number of Nash equilibria in the underlying El Farol game is countable. Therefore, the joint dynamic system (1.14) possesses a finite number of fixed points points. ■

Just to be absolutely clear, the fixed points of the joint dynamic system (1.14) consist of the following:

- **Pure strategy Nash equilibria**

  These are the pure states of the joint dynamic system (1.14) that correspond to the pure strategy Nash equilibria of the underlying game.

- **Symmetric mixed strategy Nash equilibrium**

  This is the full interior state of the joint dynamic system (1.14) that corresponds to the symmetric mixed strategy Nash equilibria of the underlying game. That is, the Nash equilibrium where all players play a unique mixed strategy best response.

- **Asymmetric mixed strategy Nash equilibria**

These are boundary states of the joint dynamic system (1.14) that correspond to asymmetric mixed strategy Nash equilibria of the underlying game. By boundary states we mean those where a subset of the N players play a unique mixed strategy best response while the remainder play a pure strategy.

- **Fixed points that are not Nash equilibria**

Not all fixed points of the joint dynamic system (1.14) correspond to Nash equilibria of the underlying game. There are pure states of the joint dynamic system (1.14) that do not correspond to pure strategy Nash equilibria of the underlying game. Note that it is not possible to have interior fixed points or fixed points on some boundary of the state space of the joint dynamic system (1.14) that do not correspond to Nash equilibria of the underlying game.

Positive Convergence Result

Proposition 1.6 *The discrete time stochastic process (1.7) converges with probability one to one of the fixed points of the standard replicator dynamics (1.10).*

Proof By Lemma 1.2 the joint dynamic system (1.14) admits a strict Lyapunov function. By Lemma 1.3 the joint dynamic system (1.14) possesses a finite number of fixed points which are identical to those of the standard replicator dynamics (1.10). Therefore, by Theorem 1.2, Benaïm (1999, Corollary 6.6), the discrete time stochastic process (1.7) converges to one of the fixed points of the standard replicator dynamics (1.10). ■

1.4.2 Proof of Main Result: Second Stage

In the second part of the proof of the main result, we show that the discrete time stochastic process (1.7) does not converge to any equilibria corresponding to Nash equilibria of the underlying game which are unstable under the adjusted replicator dynamics (1.12) or equilibria that do not corresponding to a Nash of the underlying game. We tackle this in two steps.

First, we show that the stability properties of a fixed point of the joint dynamic system (1.14) are entirely determined by the stability properties of the corresponding fixed point under the adjusted replicator dynamics (1.12). We then determine the stability properties of the Nash equilibria under the adjusted replicator dynamics (1.12). We conclude that only the pure strategy Nash equilibria are stable under the adjusted replicator dynamics (1.12). Finally, we employ Hopkins and Posch (2005, Proposition 2) to show that the discrete time stochastic process (1.7) cannot converge to any fixed points unstable under the adjusted replicator dynamics (1.12).

Second, we employ Hopkins and Posch (2005, Proposition 3) to demonstrate that the discrete time stochastic process (1.7) cannot converge to any fixed point not corresponding to a Nash equilibria under the underlying game. Therefore, we have our negative convergence result.

Unstable Equilibria in the Adjusted Replicator Dynamics

Definition 1.9 *A fixed point $\bar{x} = (\bar{y}, \bar{\mu})$ of the joint dynamic system (1.14) is unstable if its linearisation evaluated at $\bar{x}$ has at least one eigenvalue with a positive real part.*

Theorem 1.4 (Hopkins and Posch (2005, Proposition 2)) *Let $\bar{x}$ be a Nash equilibrium that is linearly unstable under the adjusted replicator dynamics (1.12). Then the Erev and Roth (1998) reinforcement learning process defined by the choice rule (1.5) and the updating rule (1.6) asymptotically converges to one of these points with probability zero.*

Lemma 1.4 *The stability properties of the fixed points of the joint dynamic system (1.14) are entirely determined by the stability properties of the corresponding fixed points of the adjusted replicator dynamics (1.12).*

Proof The linearisation of the joint dynamic system (1.14) at any fixed point, $\bar{x}$, will be of the form:

$$\begin{pmatrix} \frac{\partial \dot{y}}{\partial y} & \frac{\partial \dot{y}}{\partial \mu} \\ \frac{\partial \dot{\mu}}{\partial y} & \frac{\partial \dot{\mu}}{\partial \mu} \end{pmatrix} \tag{1.20}$$

Consider the partitions of the above matrix (1.20) evaluated at a fixed point of the joint dynamic systems (1.14) in turn. $d\dot{y}/d\mu$ is obviously the null matrix.

$$\frac{\partial \dot{y}^i}{\partial \mu^j} = 0 \text{ for all } i, j \tag{1.21}$$

Given (1.21), every eigenvalue of the matrix (1.20) is an eigenvalue for either $d\dot{y}/dy$ or $d\dot{\mu}/d\mu$. The latter matrix is diagonal.

$$\frac{\partial \dot{\mu}^i}{\partial \mu^j} = 0 \text{ for } i \neq j$$

$$\frac{\partial \dot{\mu}^i}{\partial \mu^j} \neq 0 \text{ for } i = j$$

And the diagonal elements are all negative.

$$\frac{\partial \dot{\mu}^i}{\partial \mu^i} = 1 - 2\bar{\mu}^i \left[\hat{\sigma}^i \left(0, y_t^{-i}\right) + \bar{y}^i \left[\hat{\sigma}^i \left(0, y_t^{-i}\right) - \hat{\sigma}^i \left(1, y_t^{-i}\right)\right]\right] < 0$$

Therefore, all the eigenvalues of $d\dot{\mu}/d\mu$ are negative. Now consider the elements of $d\dot{y}/dy$. This is the linearisation, or otherwise referred to as the Jacobian, of the adjusted replicator dynamics (1.12).

$$J = \begin{pmatrix} \frac{\partial \dot{y}^1}{\partial y^1} & \frac{\partial \dot{y}^1}{\partial y^2} & \cdots & \frac{\partial \dot{y}^1}{\partial y^N} \\ \frac{\partial \dot{y}^2}{\partial y^1} & \frac{\partial \dot{y}^2}{\partial y^1} & \cdots & \frac{\partial \dot{y}^2}{\partial y^N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \dot{y}^N}{\partial y^1} & \frac{\partial \dot{y}^N}{\partial y^2} & \cdots & \frac{\partial \dot{y}^N}{\partial y^N} \end{pmatrix}$$

If the linearisation of the adjusted replicator dynamics (1.12) has one or more positive eigenvalues, then the fixed point of the joint dynamic system (1.14) at which the

Jacobian is evaluated is unstable for the joint dynamic system (1.14). Otherwise, the fixed point is asymptotically stable for the joint dynamic system (1.14). ■

Now consider the stability properties of the fixed points of the adjusted replicator dynamics (1.12) that correspond to Nash equilibria in the El Farol game.

Lemma 1.5 *The fixed points of the adjusted replicator dynamics (1.12) corresponding to the pure strategy Nash equilibria of the El Farol stage game are asymptotically stable.*

Proof Given that the pure strategy Nash equilibria are strict, they constitute an evolutionary stable strategy of the El Farol stage game. By Weibull (1995), all evolutionary stable strategies are asymptotically stable under the replicator dynamics.

■

Lemma 1.6 *The fixed point of the adjusted replicator dynamics (1.12) corresponding to the symmetric mixed strategy Nash equilibrium of the El Farol stage game is asymptotically unstable.*

Proof The fixed point of the joint dynamic system (1.14) corresponding to the symmetric mixed strategy Nash equilibria of the El Farol stage game is unique and is a fully mixed equilibrium. Furthermore, at this fully mixed fixed point of the joint dynamic system (1.14), $\bar{y}^i = \bar{y}^j$. Let us consider the diagonal elements of J:

$$\frac{\partial \dot{y}^i}{\partial y^i} = \mu^i \left[1 - 2y^i\right] \left[\hat{\sigma}^i \left(1, y_t^{-i}\right) - \hat{\sigma}^i \left(0, y_t^{-i}\right)\right]$$
$$= 0 \text{ if } y^i = \bar{y}$$

Since all the diagonal elements of J equal zero, the trace of J is zero. Now consider the off diagonal elements:

$$\frac{\partial \dot{y}^i}{\partial y^j} = y^i \left[1 - y^i\right] \left[\begin{array}{c} \mu^i \left(\frac{\partial \left[\hat{\sigma}^i \left(1, y_t^{-i}\right)\right]}{\partial y^j} - \frac{\partial \left[\hat{\sigma}^i \left(0, y_t^{-i}\right)\right]}{\partial y^j} \right) \\ + \frac{\partial \mu^i}{\partial y^j} \left[\hat{\sigma}_t^i \left(1, y_t^{-i}\right) - \hat{\sigma}_t^i \left(0, y_t^{-i}\right)\right] \end{array} \right]$$

Since all players earn the same payoff in this fully mixed symmetric equilibrium, we have that $\mu^i = \mu^j$ and therefore, J is symmetric. Therefore, J has no complex eigenvalues. With a zero trace, the real eigenvalues sum to zero. Therefore, there must be at least one eigenvalue which is positive. Hence, $\bar{x}$ is linearly unstable with respect to the joint dynamic system (1.14). ■

Lemma 1.7 *The fixed points of the adjusted replicator dynamics (1.12) corresponding to the asymmetric mixed strategy Nash equilibria of the El Farol stage game are asymptotically unstable.*

Proof At the fixed points of the joint dynamic system (1.14) corresponding to the asymmetric mixed strategy Nash equilibria, $N - j - k$ players randomise over entry while the remaining $j + k$ players play a pure strategy. One can then calculate the Jacobian, J, evaluated at this fixed point which is of the form:

$$J = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

where A is a $(N - j - k) \times (N - j - k)$ matrix of the form found at the symmetric fixed point as described in Lemma 1.6.

It is easily verified that C is a diagonal matrix of negative elements. By the same argument as put forward in Lemma 1.6, A is a mixture of positive and negative eigenvalues. Therefore, J has at least one positive eigenvalue and it follows that the fixed point associated with the asymmetric mixed strategy Nash equilibrium is unstable under the adjusted replicator dynamics (1.12). ■

Non-Nash Fixed Points of the Joint Dynamic System

Theorem 1.5 (Hopkins and Posch (2005, Proposition 3)) *Let $\bar{x}$ be a fixed point of the replicator dynamics (1.10) which is not a Nash equilibrium. Then the Erev and Roth (1998) reinforcement learning process defined by choice rule (1.5) and the updating rule (1.6) asymptotically converges to one of these points with probability zero.*

Therefore, the discrete time stochastic process (1.7) cannot converge to any fixed point not corresponding to a Nash equilibrium under the underlying game.

Negative Convergence Result

Proposition 1.7 *The discrete time stochastic process (1.7) converges with probability zero to equilibria corresponding to Nash equilibria of the underlying game unstable under the adjusted replicator dynamics (1.12) or equilibria not corresponding to a Nash equilibrium of the underlying game.*

Proof The result follows from Theorem 1.4, Hopkins and Posch (2005, Proposition 2), and Theorem 1.5, Hopkins and Posch (2005, Proposition 3). ∎

1.4.3 Proof of Main Result: Concluding Stage

Proposition 1.8 *In the El Farol game with identical bounded rational agents learning in accordance with the Erev and Roth (1998) reinforcement learning model, long-run behaviour converges asymptotically to the set of pure strategy Nash equilibria of the El Farol stage game.*

Proof The result follows directly from our positive convergence result, Proposition 1.6, and our negative convergence result, Proposition 1.7. ∎

1.5 Conclusion

The results obtained from modelling the El Farol bar problem as a repeated game with boundedly rational agents implies that people tend to minimise bad experiences and maximise good ones. This is exactly what is assumed by the Erev and Roth (1998) reinforcement learning model.

The application of the Erev and Roth (1998) reinforcement learning model implies that the average attendance converges to the capacity of the El Farol bar as in Arthur's (1994) 'inductive thinking' approach to modelling boundedly rational agents. The difference lies in who, in the long-run, attends the bar. The most salient aspect of this result is that in the El Farol game the population of boundedly rational agents, who behave in accordance with the Erev and Roth (1998) reinforcement learning model, are partitioned into those who always attend the El Farol bar and those who always stay at home. This differs from the outcome of Arthur's (1994) model where agents are differentiated by the forecasting methods, not by attendance.

The main result implies sorting and is crucially dependent on the fact that the game in question is a potential game. It can be shown that the result is robust when compared to other variants of reinforcement learning. In fact Duffy and Hopkins's (2005) paper shows how this would be the case with stochastic fictitious play. It is also possible to derive some general results for an extension to this treatment where players have idiosyncratic payoff functions. Milcataich (1996) shows that any multiplayer coordination game with two identical pure actions for each player admits a potential function and by definition, is a potential game. Therefore, even if one considers players in the El Farol game with heterogenous preferences, it appears that reinforcement learning will lead to sorting.

Appendix 1.A

1.A.1 Proof to Proposition 1.1

Proof The number of pure strategy Nash equilibria in the one-shot El Farol game is the number of ways C different players can be chosen out of the set of N players at a time. ■

1.A.2 Proof to Proposition 1.2

Proof *Existence*

Define the binary state of the bar, either uncrowded or crowded, as a binomial distribution, denoted $P^i(N, C, \alpha)$, over the number of players in the game, denoted by N, the capacity of the bar, denoted by C, and the mixed strategies employed by player i, denoted by the probability α of attending and $(1-\alpha)$ of staying at home.

In particular in a mixed strategy equilibria, each player i should be indifferent between going to the bar and staying at home.

$$E\left[\pi^i|\delta=1\right] = E\left[\pi^i|\delta=0\right], \qquad \forall \quad i \in \{1, 2, ..., N\}$$

We have that:

$$\begin{aligned} &E\left[\pi^i|\delta=1\right] = G \cdot P^i(N, C, \alpha) + B \cdot \left[1 - P^i(N, C, \alpha)\right] \\ &E\left[\pi^i|\delta=0\right] = S \end{aligned} \tag{1.22}$$

where $P^i(N, C, \alpha)$ denotes the probability of the bar being in the uncrowded state or otherwise stated as the probability that less than C other players out of the $(N-1)$ remaining players choose to attend the bar.

Given that players are homogeneous in preferences and the El Farol game is symmetric in payoffs, we may write:

$$P^i(N, C, \alpha) = P(N, C, \alpha) \text{ for all } i \in \{1, 2, ..., N\}.$$

Returning to (1.22), we can now substitute $P(N, C, \alpha)$ for $P^i(N, C, \alpha)$ and solve for $P(N, C, \alpha)$.

$$P(N, C, \alpha) = \frac{(S - B)}{(G - B)}$$

where $P(N, C, \alpha)$ is defined by the following binomial probability:

$$P(N, C, \alpha) = \sum_{m=0}^{C-1} C_m^{N-1} \alpha^m (1 - \alpha)^{(N-1)-m}.$$

Therefore, we have (1.3).

Uniqueness:

$P(N, C, \alpha)$ is continuous and well defined over the closed interval $[0, 1]$. Given that $\lim_{\alpha \to 0} P(N, C, \alpha) = 1$ and $\lim_{\alpha \to 1} P(N, C, \alpha) = 0$, (1.3) has a unique solution if the partial derivative of $P(N,C,\alpha)$, denoted $P_\alpha(N,C,\alpha)$, is less than zero.

$$\begin{aligned}
P_\alpha(N,C,\alpha) &= \frac{\partial}{\partial\alpha}\left(\sum_{m=0}^{C-1} C_m^{N-1}\alpha^m(1-\alpha)^{(N-1)-m}\right) \\
&= \left(\sum_{m=1}^{C-1}(N-1)\,C_{m-1}^{N-2}\alpha^{m-1}(1-\alpha)^{(N-1)-m}\right) \\
&\quad -\left(\sum_{m=0}^{C-1}(N-1)\,C_m^{N-2}\alpha^m(1-\alpha)^{((N-1)-m)-1}\right)
\end{aligned}$$

Now let $k=(m-1)$.

$$\begin{aligned}
P_\alpha(N,C,\alpha) &= (N-1)\left(\sum_{k=0}^{C-2} C_k^{N-2}\alpha^k(1-\alpha)^{N-2-k}\right) \\
&\quad -(N-1)\left(\sum_{m=0}^{C-1} C_m^{N-2}\alpha^m(1-\alpha)^{N-2-m}\right) \\
&= (N-1)(-1)\,C_{m-1}^{N-2}\alpha^{m-1}(1-\alpha)^{N-1-m} \\
&< 0
\end{aligned}$$

Hence (1.3) has a unique solution and thus the number of mixed strategy Nash equilibria where all agents employ the same 'mixing' is one. ■

1.A.3 Proof to Proposition 1.3

Proof Here we use the fact that there are only two pure strategies available to each player. We show via contradiction that if no players employ a pure strategy, all players must play the same mixed strategy.

First, assume that here are two players, say i and j, who employ different mixed strategies in a mixed strategy equilibrium in which their probabilities of attending the bar are α_i and α_j respectively, where $\alpha_i \neq \alpha_j$. Note that the remaining $N-2$ players also play mixed strategies. Since player i uses a mixed strategy in equilibrium, it must be that the probability of less than C attending the bar is $(S-B)/(G-B)$ given α_j and the mixed strategies employed by the $(N-2)$ remaining players. Likewise if player j uses a mixed strategy in equilibrium, it must be that the probability of less than C attending the bar is $(S-B)/(G-B)$ given α_i and the mixed strategies employed by the $(N-2)$ remaining players. If $\alpha_i \neq \alpha_j$, both these statements cannot be true.

Consider the case where $\alpha_i > \alpha_j$. The probability of less than C attending the bar is $(S-B)/(G-B)$ given α_j and the mixed strategies employed by the $(N-2)$ remaining players. It is then impossible to have the probability of less than C attending the bar, given α_i and the mixed strategies employed by the $(N-2)$ remaining players equal $(S-B)/(G-B)$. We have the similar argument for $\alpha_i < \alpha_j$. Recall that agents have only two possible pure strategies.

This contradiction tells us that in an equilibrium where all players play a mixed strategy, they must all play the same mixed strategy. ∎

1.A.4 Proof to Proposition 1.4

Proof Recall that an asymmetric mixed strategy Nash equilibria is a Nash equilibrium where players from a subset of the population play either of the available pure strategies, and the remaining players play the symmetric mixed strategy which supports the asymmetric mixed strategy Nash equilibria. In an asymmetric mixed strategy equilibrium, we require at least two players to play a mixed strategy and at least one player to play a pure strategy. Given that all the players playing a mixed strategy are playing the same mixed strategy, we can simply count the number of asymmetric mixed strategy Nash equilibria. ■

Appendix 1.B

1.B.1 Proof to Proposition 1.5

Proof First, suppose that in period t player i chooses to attend the bar. The ith player's strategy adjustment of going to the bar given that player i chooses to go to the bar is,

$$\begin{aligned}
E\left[y_{t+1}^i | \delta_t^i = 1\right] - y_t^i &= \frac{q_t^i(1) + \hat{\sigma}_t^i\left(1, y_t^{-i}\right)}{Q_t^i + \hat{\sigma}_t^i\left(1, y_t^{-i}\right)} - \frac{q_t^i(1)}{Q_t^i} \\
&= \frac{\left[1 - y_t^i\right]\hat{\sigma}_t^i\left(1, y_t^{-i}\right)}{Q_t^i} \\
&\quad - \left[\left[1 - y_t^i\right]\left[\frac{\left(\hat{\sigma}_t^i\left(1, y_t^{-i}\right)\right)^2}{Q_t^i\left(Q_t^i + \hat{\sigma}_t^i\left(1, y_t^{-i}\right)\right)}\right]\right] \\
&= \frac{\left[1 - y_t^i\right]\hat{\sigma}_t^i\left(1, y_t^{-i}\right)}{Q_t^i} + O\left(\frac{1}{\left(Q_t^i\right)^2}\right)
\end{aligned}$$

Similarly, the ith player's strategy adjustment of going to the bar given that player i chooses to stay at home is,

$$\begin{aligned}
E\left[y_{t+1}^i | \delta_t^i = 0\right] - y_t^i &= \frac{q_t^i(1)}{Q_t^i + \hat{\sigma}_t^i\left(0, y_t^{-i}\right)} - \frac{q_t^i(1)}{Q_t^i} \\
&= \frac{-y_t^i\hat{\sigma}_t^i\left(0, y_t^{-i}\right)}{Q_t^i} - \left[y_t^i\left[\frac{\left(\hat{\sigma}_t^i\left(0, y_t^{-i}\right)\right)^2}{Q_t^i\left(Q_t^i + \bar{\sigma}_0^i\right)}\right]\right] \\
&= \frac{-y_t^i\hat{\sigma}_t^i\left(0, y_t^{-i}\right)}{Q_t^i} + O\left(\frac{1}{\left(Q_t^i\right)^2}\right)
\end{aligned}$$

Recall that by definition player i goes to the bar with probability y_t^i and stays at home with probability $\left[1 - y_t^i\right]$. Therefore, the expected motion of the ith player's

$\delta = 1$ strategy adjustment in the repeated El Farol game is,

$$E\left[y_{t+1}^i | y_t^i\right] - y_t^i = y_t^i \left[\frac{\left[1 - y_t^i\right] \hat{\sigma}_t^i \left(1, y_t^{-i}\right)}{Q_t^i}\right] + \left[1 - y_t^i\right] \left[\frac{-y_t^i \hat{\sigma}_t^i \left(0, y_t^{-i}\right)}{Q_t^i}\right] + O\left(\frac{1}{\left(Q_t^i\right)^2}\right)$$

and after further algebraic simplification we arrive at (1.7). ■

Chapter 2

Why People Don't Play Mixed Strategies: Learning in Finite Population Games

2.1 Introduction

Randomisation is central to mixed strategies. One interpretation as to how individuals implement mixed strategies is that they make their choices based on a random lottery. However, this view has its critics as people are generally considered to be poor at generating random outcomes. An alternative interpretation of mixed strategies, imagines that players represent a population of agents. Each of the agents chooses a pure strategy and the payoff depends on the fraction of agents choosing each strategy. The mixed strategy then represents the distribution of pure strategies chosen by each population.

The issues of interpretation becomes particularly problematic in games with only mixed strategy equilibria. In tackling this question we investigate games with mixed strategy equilibria and random matching. The addition of random matching allows for both interpretations of mixed strategies to coexist. Random matching allows for the second interpretation of mixing, but equally each agent could still choose to randomise their individual choices. Therefore, they have the choice as to how the 'mixing' is implemented.

When modelling games with mixed strategy equilibria and random matching, there are effectively two game interpretations: the two-player game individuals play once they are matched and the larger game, which involves a population of players matched into pairs, all of whom play the two player game simultaneously.

In this paper we study the long-run behaviour of a single finite population of boundedly rational agents learning in accordance with a reinforcement learning model, randomly matched to play a symmetric two-player game. Our model comprises four component elements: the population setting, the matching technology, the assumed individual learning model, and the two-player games played once individuals are matched into pairs. Our overall objective is to determine, using the tools of learning theory, the long-run convergence of play in the finite population game as defined below. We find that, for a large class of games that only have mixed equilibria, reinforcement learning predicts that, under random matching, individual behaviour will converge to play of fixed pure strategies.[1]

Our contribution to the literature on population games and the theory of learning is fourfold. First, we apply the Erev and Roth (1998) model of reinforcement learning to finite population games with random matching. We believe that there are genuine contributions to be made in this area. Once the long-run behaviour convergence results are proved for individuals learning in accordance with the Erev and Roth (1998) model of reinforcement learning, these results can be easily extended to other models of individual learning. We provide much of the structure to prove analytically that long-run behaviour will converge asymptotically to the set of pure strategy Nash equilibria of the one-shot finite population game corresponding to the mixed strategy Nash equilibrium in the population game. In other words, in aggregate, the

[1]On reflection this result should not be surprising, as we would expect convergence to this set of equilibria in games with strict pure strategy equilibria.

distribution of pure strategy players in the finite population game will correspond to the evolutionary stable equilibrium in the corresponding population game. Our work suggests that in finite anti-coordination population games, learning theory predicts sorting. In a single finite population of players who are learning in accordance with the Erev and Roth (1998) model of reinforcement learning, theory suggests that the population will eventually be subdivided into distinct groups; each group specialising and invariably choosing to play one pure strategy action.

Second, we provide an alternative interpretation of results from experiments of population games with random matching. Friedman (1996), amongst others, has carried out a number of laboratory experiments in economics to test predictions of evolutionary game theory. One of the most consistently tested population games in laboratory environments is the Hawk-Dove game. Evolutionary game theory predicts that play in the population should converge to the symmetric mixed strategy Nash equilibrium of the Hawk-Dove game. Interestingly, replicator dynamics, as introduced by Maynard Smith (1974), assumed that all individuals are hard wired to play a specific type of strategy, and it is the frequency of strategy types that evolves in the population. Crucially, this analysis assumes an infinite population pool from which players are drawn. However, in a laboratory setting this is not a practical or easily incorporated design feature. Therefore, most of the predictions of evolutionary game theory on population games are actually tested on finite populations. Although the experimental evidence broadly supports the predictions of evolutionary game theory on aggregate, there are notable departures once individual behaviour is examined. There is consistent evidence to suggest that, on an individual level, players do seem to favour one pure strategy or another. Friedman (1996) suggests that Harsanyi's purification approach might provide an explanation for these observations (Harsanyi and Selten 1988). Our work suggests an alternative interpretation as to why some players usually play Hawk and others Dove; namely the finite population aspect of the experimental design.

Third, we wish to highlight the importance of random matching in population games. Our results appears to indicate that random matching actually takes on the role of randomisation. Instead of individuals randomising their choices, such as playing mixed strategy profiles, the random matching effectively takes on this task that would be left to individuals in the same two-player games with fixed matching. We believe this is an important observation, as it indicates that the appropriate modelling of matching protocols in any theoretical or applied economic application has significant effects on predictions of models, and therefore, provides lessons for policy makers.

Finally, we wish to bring to the attention of economic theorists and experimental researchers, existing literature from the field of complex systems; namely, Borkar, Jain, and Rangarajan (1998), Borkar, Jain, and Rangarajan (1999) and Borkar, Jain, and Rangarajan (2002). It is a well-known phenomenon that a wide variety of complex adaptive systems exhibit, on the micro or component level, a high degree of specialisation, while maintaining, in a macro sense, ever-increasing diversity. In our work we make progress in demonstrating that in finite population games with random matching, composed of two-player games with opposing interests, the theory of individual learning suggests that environments will naturally evolve to exhibit both greater diversity and specialisation.

The intuition behind our results is that the matching protocol and population setting can transform a strategic game with interior evolutionary stable strategies to one with strict pure strategies which have been shown in the literature to be evolutionarily stable (see Weibull 1995).

In Section 2.2 we review the finite population game framework. We define the finite population game and discuss the four defining components of our model framework in detail. We also suggest particular finite population games that exhibit properties that are attractive for further study. In section 2.3 we state and develop a proof of our main result; that is, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a strict pure Nash equilibrium of the one-shot finite population game with random matching. In Section 2.4 we present evidence from computer simulations supporting out main result. Finally, in Section 2.5 we provide some concluding remarks.

2.2 Finite Population Games

In this paper we study the long-run behaviour of a single finite population of boundedly rational agents learning in accordance with a reinforcement-learning model, randomly matched to play a symmetric two-player game. Our model shall consist of four component elements: the population setting, the matching technology, the assumed individual learning model, and the two-player games played once individuals are matched into pairs. Our overall objective is to determine, using the tools of learning theory, the long-run convergence of play in the finite population game as defined below.

Definition 2.1 *Define the* $S \times S$ *finite population game with random matching as the infinitely repeated game where in each period all individuals from a finite population* $N \in 2\mathbb{N}$ *are randomly matched into pairs to play the one-shot symmetric two-player game* $\Gamma =< \Delta, \pi >$ *consisting of,*

- *a finite and identical set of actions for each player, denoted by* Δ *and indexed by* δ
- *a payoff function* $\pi^i : \delta^i \times \delta^{-i} \rightarrow \mathbb{R}$

where δ^i *denotes the action taken by player* i *and* δ^{-i} *denotes the action taken by player* i*'s opponent.*

Let us begin by discussing the components of this model framework in further detail, starting with the population setting.

2.2.1 Population Setting

In principle there are two population settings that researchers have considered in this type of model framework, infinite and finite. The use of the infinite population is the more traditional assumption and the more widely used for two reasons. The first is tractability and the second is that in many economic applications this might be a reasonable assumption to make.

The paramount implication of assuming that the population is infinite is that economists are really assuming that individuals perceive their actions as having no effect on the average play or frequencies in the larger population game. This is a very important simplification for two reasons. First, it allows for tractability in understanding the overall dynamics in the population game. Second, it removes strategic motivations from the level of the individual agents. In population games with infinite populations, researchers are generally interested in the evolution of the frequency of type of player in the population over time. In essence individuals in the infinite population are hard-wired to be of a certain type, and it is the evolution of their frequency, i.e. the frequency of each type of player in the population, that researchers are interested in studying.

In this study we do not make this infinite population assumption. Specifically, we assume that the number of individuals in the population is finite. Furthermore, the population size is small enough so that individual players realise that there is a difference between the overall average play in the population and the average play of those other individuals they are likely to be matched with to play the two-player game. Obviously, if the finite population is very large, it is reasonable to assume that individuals would treat these two different measures as the same. It is for this reason that the infinite population assumption is justified in most economic applications.

The salient aspect of this component of the population game model is that individuals are drawn from a common pool, or single finite population, and matched pair wise, according to a pre-specified matching technology, to play the two-player game. One of the contributions of this paper is that we demonstrate that the long-run outcomes of players drawn from finite populations are both qualitatively and quantitatively different from those outcomes where individuals are drawn and matched into pairs from an infinite population.

In particular this proves to be the case for the much studied Hawk-Dove game which is discussed in detail later in this paper. Standard results from evolutionary game theory prove that the interior mixed strategy equilibrium is an evolutionarily stable strategy (ESS). In our analysis, assuming a finite population, we show that while this might be the case in aggregate, i.e. the frequency of types is interior in equilibrium, on an individual level everyone in the population will actually be playing a pure strategy and the population game will evolve to a pure strategy Nash state over time.

An interesting issue to address at this point is the general interpretation of mixed strategies in population games (Hofbauer 2000). It is true that in the population game framework, as devised by Maynard Smith (1974), it is effectively assumed that individual players are playing pure strategies. In the work presented here, we arrive at the same outcome at the individual level, but we do so without making any assumption as to the type of player each individual in the population is. It is the result of the ecology of the dynamic system that all players in equilibrium will play pure strategies; all individuals will, in the end, be 'type cast' in equilibrium. This result arises without any preemptive hard wiring of individual's strategies.

There is one final matter with regard to our assumptions over the population setting that requires explanation. Given that in our model, all players are matched into pairs in each period, it is convenient to assume that the finite population size, N, is an element of the even integers. If this were not the case, certain matching technologies might leave some individuals unmatched in each round. This would add unnecessary complication to the application of our learning model. It is felt that leaving some players unmatched would add few further insights to the research and it would be best to maintain some flexibility as to which matching technology is assumed. Different matching schemes could possibly offer advantages or disadvantages in economic applications and experimental settings. This brings us on to the next component of our framework, the matching technology.

2.2.2 Matching Technology

The principle matching technology that we shall consider throughout this paper is random matching. This is a commonplace assumption in the literature on learning and evolution. However, as pointed out by Hopkins (1999) there are several ways of modelling this type of interaction.

Random Matching

Fudenberg and Kreps (1993) proposed three alterative random matching schemes.

Scheme 1 At each date t, one pair of players is selected to play a one-shot two-player game. Once the game is played, their actions are revealed to all potential players. Those who played at date t are then returned to the pool of potential players.

Scheme 2 At each date t there is a random matching of all players, so that each player is paired with another player with whom the game is played. At the end of the period, it is reported to all players how the entire population played. The play of any particular player is never revealed.

Scheme 3 At each date t there is a random matching of the players into pairs and each pair of players then plays the game. Each player recalls at date t what happened in the previous encounters in which he was involved, without knowing anything about the identity or experience of his rivals.

It is worth pointing out that the scheme assumed as the basis of replicator dynamics is Scheme 3. Boylan (1992) has shown that when a population is assumed to be infinite, the dynamics are deterministic. The beauty of this scheme is that it is decentralised in the sense that it does not, as opposed to Scheme 1 or 2, require any public announcements. However, Hopkins (1999) notes that there are other matching technologies similar to Scheme 2 that which do not require any public announcements, namely:

Scheme 2a In each round, players are matched according to Scheme 1 or 3 an infinite number of times.

Scheme 2b In each round there is a "round-robin" tournament, where each player meets each of his potential opponents exactly once.

As Hopkins (1999) points out, Scheme 2a and Scheme 2b have been widely employed in the learning literature primarily for reasons of tractability. In addition they ensure a deterministic result to the matching technology even when the population is finite. Furthermore, with Schemes 2, 2a and 2b, all players know the exact distribution of strategies in the population when choosing their next strategy. This is not necessarily the case with Scheme 3. Hopkins (1999) shows that both Schemes 2 and 3 have the same continuous time limit when an infinite population is assumed.

Given these properties it is convenient to employ one of the variants of Scheme 2 in our analysis. Variants of Scheme 2 also provide a suitable framework to test any theoretical results in an experimental laboratory.[2]

2.2.3 Individual Learning Model

Now we turn our attention to the learning model we shall assume individual agents adhere to when updating their strategies from period to period. The contribution we wish to make to the literature on population games is that in finite population games with random matching, reinforcement learning predicts that long-run behaviour in the population will, in aggregate, mimic the evolutionary stable equilibrium and, on the individual level, all players will always play pure strategies.

[2]It is also worth noting that there are other possible matching schemes. In fact, it is quite likely the case that our chosen matching technology - being a variant of random matching - is not suitable or a reasonable approximation of actual interaction in some economic and social situations.

The benchmark for this type of analysis is the replicator dynamics. Indeed, we shall study the replicator dynamics for our finite population games in detail. However, as previously alluded to, this paper aims to bring some of the results from learning theory to the study of finite population games. Finite populations games often represent complex strategic environments where rational deductive thinking fails to provide any clear solution. This could be the case for a variety of reasons, including cognitive limitations and/or equilibrium refinement/coordination. We posit that in environments that are either complicated and/or ill-defined, individuals adopt inherently inductive methods or heuristics to maximise payoffs.

The theory of individual learning represents one of the frameworks that aims to model decision making where perfect rationality fails. In addition there is ample anecdotal and experimental evidence supporting the notion that humans do not always make choices using perfect rationality. Learning theory postulates that individuals find their way to an optimal solution via trial and error. Effectively, we wish to model the boundedly rational agents in our finite population games as individuals who gradually adjust their behaviour over time until there is no longer room for improvement. Therefore, these individuals find their optimal strategies through repeated interaction and the application of an adaptive algorithm.

We assume that any adaptive learning algorithm adheres to some basic principles of individual learning. First is the law of effect: choices that have led to good outcomes in the past are more likely to be chosen again in the future. Second is the power law of practice: learning curves should initially be steep and then later they should be flatter. This is paramount to assuming that in any adaptive process the adjustments become smaller over time. Finally, choice behaviour should be probabilistic. This is a basic assumption in most mathematical learning theories proposed in psychology. Erev and Roth (1998) have developed a robust model of reinforcement learning, which incorporates these principles. We shall apply this model to our finite population games with random matching.

Note that in addition to Erev and Roth's (1998) model of reinforcement learning, there are other available alternatives. Before we outline Erev and Roth's (1998) model of reinforcement learning in detail, it is worth mentioning some of the advantages of this particular individual learning model. First is that once we have established some results in finite population games assuming Erev and Roth's (1998) model of reinforcement learning, we can easily extend these results to other models of individual learning such as hypothetical reinforcement learning (also referred to as fictitious play learning). Second, there is a body of results on learning already in existence that we can apply to our research to further understand the long-run convergent behaviour of individuals in finite population games with random matching. We shall set out our methodology in detail in Section 2.3.

Erev and Roth (1998) Reinforcement Learning

We now set out the procedure for the Erev and Roth (1998) reinforcement learning model in detail. In this learning model, each player i has a propensity to undertake each action in each period, denoted $q_t^i(\delta)$. The timeline of the learning algorithm is that in each period t, each player i chooses to undertake one of their available actions $\delta \in \Delta$ in accordance with a mapping from the propensities to the unit interval $[0,1]$. This mapping is defined by the choice rule. The player i then undertakes the action dictated by the choice rule and receives a payoff in that period associated with that action. Player i then updates his propensities. The updating procedure is determined by the updating rule. In the Erev and Roth (1998) reinforcement learning model, the only propensities to be updated are those corresponding to the actual action taken.

We can now define the model formally. The learning procedure comprises of three components: the initial conditions, a choice rule and an updating rule.

Initial Conditions Let $q_t^i(\delta)$ be player i's propensity to play action $\delta \in \Delta$ in period t. In the initial period, $t = 0$, we assume that all players have positive propensities for all possible actions. That is,

$$q_t^i(\delta) > 0 \text{ for } t = 0 \text{ and for all } i \in N \text{ and } \delta \in \Delta \tag{2.1}$$

This assumption, along with positive payoffs, will also ensure that $q_t^i(\delta) > 0$ for all t and $\delta \in \Delta$.

Choice Rule Each player i has a positive propensity, $q_t^i(\delta)$, to take action $\delta \in \Delta$ in period t. In models of reinforcement learning, the choice rule provides a mapping from propensities to strategies. Let $(y_t^i, [1 - y_t^i])$ represent player i's mixed strategy in period t with two possible actions $\delta \in \Delta = \{0, 1\}$, where y_t^i is the probability placed by agent i on action $\delta = 1$ in period t and $[1 - y_t^i]$ is the probability placed by agent i on action $\delta = 0$ in period t. The choice rule employed in the Erev and Roth (1998) reinforcement learning model is often referred to as the simple choice rule. It is a straightforward probability mapping from propensities to the unit interval $[0, 1]$. That is,

$$\Pr(\delta = 1) = y_t^i = \frac{q_t^i(1)}{\sum_{\delta \in \Delta} q_t^i(\delta)} = \frac{q_t^i(1)}{Q_t^i} \tag{2.2}$$

where $Q_t^i = \sum_{\delta \in \Delta} q_t^i(\delta)$.[3]

Updating Rule Let $\bar{\sigma}^i(\delta_t^i, m_t^{-i})$ denote the realised increment to player i's propensity in period t from taking action $\delta \in \Delta$ given the aggregate actions taken by the remaining $N - 1$, denoted by m_t^{-i} where $m_t^{-i} = \sum_{j \neq i} \delta_t^j$. To complete, and most crucial to, our reinforcement learning model, we must state the means by which players update their propensities. Specifically, in the Erev and Roth (1998) reinforcement learning model, if agent i takes action δ in period t, then the agent's

[3]Note that when there are only two possible actions for each player i we can write

$$\Pr(\delta = 0) = (1 - y_t^i) = \frac{q_t^i(0)}{\sum_{\delta \in \Delta} q_t^i(\delta)} = \frac{q_t^i(0)}{Q_t^i}$$

δth propensity is increased by an increment equal to agent i's realised payoff in that period. All other propensities remain unchanged. In other words only realised payoffs act as reinforcers. We thus have the following updating rule,[4]

$$q_{t+1}^{i}(\delta) = q_{t}^{i}(\delta) + \bar{\sigma}^{i}\left(\delta_{t}^{i}, m_{t}^{-i}\right) \text{ for all } \delta \in \Delta = \{0, 1\} \tag{2.3}$$

2.2.4 Pair Game

The final component to our finite population game is the two-player game individuals play once they are matched. To avoid any confusion we shall refer to this game as the pair game.

The class of two-player games consists of both symmetric and asymmetric games. Furthermore, there are two-player games where players have access to a finite or infinite set of strategies. We shall only consider two-player games where players have a finite set of actions available. We begin our analysis with definitions of symmetric and doubly symmetric two-player games.

Definition 2.2 *A two-player game is a symmetric game if both players have the same strategy set and the second player's payoff matrix, B, is the transpose of the*

[4]Note that this updating rule reveals why in this model of reinforcement learning all payoffs must be positive. Otherwise, there would be a possibility of propensities becoming negative and thus leading to choice probabilities that are undefined.

first player's, A; that is,

$$B = A^T \tag{2.4}$$

As Weibull (1995) points out, the requirement that the second player's payoff matrix be the transpose of the first player's is equivalent to the symmetry requirement in pure strategy payoffs for a two-player game $\Gamma = \langle N, S, \pi \rangle$ where $S_1 = S_2$; that is,

$$\pi_1 (s_2, s_1) = \pi_2 (s_1, s_2) \text{ for all } (s_1, s_2) \in S \tag{2.5}$$

Definition 2.3 *A symmetric two-player game is doubly symmetric if* $A = A^T$.

We first consider symmetric finite strategy two-player games. We study symmetric games because the analysis of asymmetric games in our finite population game framework adds complications due to the different roles that exist, i.e. the row and column players have different payoff matrices. Furthermore, with asymmetric games additional technical difficulties arise in the matching schemes. Given that players are assumed to be drawn from a single finite population, the addition of separate roles would require the matching technology to allocate these roles, i.e. whether a player is a row or a column player, as well as randomly matching individuals in the population.

Incorporating asymmetric games into our framework would be advantageous in deriving results applicable to all finite population game, but it is our belief that taking this approach would cause a pivotal aspect of the finite population setting to be lost; namely, that individual players can discount their own play from the information about average play of the entire population that is available to all players. In general in order to be able to model asymmetric finite population games, it is necessary to assume a multi-population set-up. In this approach one assumes that players fulfilling different roles are drawn from separate populations. Therefore, the average play of the population players that is relevant to, say, a row player, is the average play of the entire population of column players. Note that the row player is not a member of the population of column players.

Second, we focus on games with finite action sets for reasons of tractability, and because it is these games that have been examined in detail in experimental settings. Recall that one of the objectives of this research is to contribute to the literature on population games by showing that experimental studies designed to test theoretical findings of evolutionary game theory on populations games have, for practical reasons, done so with finite populations. We wish to provide an additional justification to the experimental evidence that suggests individual players seem to play pure strategies in these experimental designs.

We will begin by considering the full class of symmetric 2×2 games, followed by consideration of a subset of the class of symmetric $S \times S$ anti-coordination games. Finally, we study in further detail a sub-class of symmetric 3×3 anti-coordination games.

2×2 Symmetric Pair Games

Consider a generic symmetric 2×2 game with the following payoff matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.6}$$

We can normalise matrix A by subtracting a_{11} from the first column and a_{22} from the second column, arriving at the following dynamically equivalent payoff matrix:

$$A' = \begin{pmatrix} 0 & a_1 \\ a_2 & 0 \end{pmatrix} \tag{2.7}$$

The first thing to note is that the new matrix is symmetric. Therefore, and according to Definition 2.3, we have a doubly symmetric 2×2 game with the payoff matrix A'. It follows directly that any symmetric 2×2 game can be represented by a point $a = (a_1, a_2) \in \mathbb{R}^2$ on the plane. Therefore, we can easily categorise all symmetric 2×2 games into one of the following three categories (See Weibull 1995):[5]

[5]The cases where one or more inequalities are actually equalities are not considered, as they are non-generic. As pointed out by (Borkar, Jain, and Rangarajan 1998), these cases are not difficult to handle in the subsequent analysis.

Dominant Strategy Games This corresponds to symmetric 2×2 games in the northeast and southwest quadrant of the $a_1 \times a_2$ plane, i.e. where $a_1 a_2 < 0$. It is evident that in any symmetric 2×2 game of this type, strategy two strictly dominates strategy one or vice versa. Therefore, all such games are strictly dominance solvable and thus there exists a unique Nash equilibrium in pure strategies in the pair game. The prototype example of a dominant strategy symmetric 2×2 game is the Prisoner's Dilemma Game.

Example 2.1 *Prisoner's Dilemma Game.*

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$4,4$	$0,5$
	Defect	$5,0$	$3,3$

The Prisoner's Dilemma Game is characterised by the dominant strategy of defect for both players. Thus, there exists a unique Nash equilibrium in pure strategies, namely (Defect, Defect).

Coordination Games This corresponds to symmetric 2×2 games in the southeast quadrant of the $a_1 \times a_2$ plane, i.e. where $a_1 < 0$ and $a_2 < 0$. All such games in this category have two symmetric pure strategy Nash equilibria and one asymmetric mixed strategy Nash equilibrium where players play the mixed strategy $\left(\frac{a_2}{a_1+a_2}, \frac{a_1}{a_1+a_2}\right)$,

$\left(\frac{a_1}{a_1+a_2}, \frac{a_2}{a_1+a_2}\right)$ respectively. A well known example of this category of symmetric 2×2 game is the Stag-Hunt Game.

Example 2.2 *Stag-Hunt Game*

		Hunter 2	
		Stag	Rabbit
Hunter 1	Stag	$5, 5$	$0, 4$
	Rabbit	$4, 0$	$3, 3$

The Nash equilibria of the Stag-Hunt Game are (*Stag,Stag*), (*Rabbit,Rabbit*), $\left(\left(\frac{3}{4}\textit{Stag}, \frac{1}{4}\textit{Rabbit}\right), \left(\frac{1}{4}\textit{Stag}, \frac{3}{4}\textit{Rabbit}\right)\right)$.

Anti-Coordination Games This corresponds to symmetric 2×2 games in the northwest quadrant of the $a_1 \times a_2$ plane, i.e. $a_1 > 0$ and $a_2 > 0$. Again in this category no strategy is dominated. However, here a player's best reply to a pure strategy is to play the other pure strategy. Therefore, these games have two asymmetric strict Nash equilibria and one symmetric mixed-strategy Nash equilibrium. The prototype example in this category is the Hawk-Dove Game.

Example 2.3 *Hawk-Dove Game*[6]

Player 2

		Fight	Yield
Player 1	Fight	$\frac{(v-c)}{2}, \frac{(v-c)}{2}$	$v, 0$
	Yield	$0, v$	$\frac{v}{2}, \frac{v}{2}$

The Nash equilibria of the Hawk-Dove Game are $(Fight, Yield)$, $(Yield, Fight)$, $\left(\left(\frac{v}{c} Fight, \frac{c-v}{c} Yield\right), \left(\frac{v}{c} Fight, \frac{c-v}{c} Yield\right)\right)$.

Remark It is this final category of symmetric 2×2 games that is predominantly of interest. Using the standard analysis of evolutionary game theory, both dominant strategy and coordination games have pure strategy evolutionary stable strategies (ESS). In the 2×2 anti-coordination games, the ESS is strictly interior. In general we shall now refer to anti-coordination games as those for which the following anti-coordination condition holds:

Condition 2.1 *The symmetric two player pair game with payoff matrix*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{pmatrix} \tag{2.8}$$

[6]We presume that the cost of a fight exceeds the value of victory: $c > v$.

is said to be an anti-coordination game if and only if it is diagonally subdominant. That is, if

$$a_{\alpha\alpha} < a_{\alpha\beta} \text{ for all } \alpha \neq \beta \text{ and } \alpha, \beta \in S \tag{2.9}$$

There are two things worth noting about 2×2 symmetric pair games. The first is that, as we have already addressed, all 2×2 symmetric pair games are in fact doubly symmetric. This is a very important property that we shall make use of to show that the finite population game with random matching, consisting of a 2×2 symmetric pair game, admits a potential function and is, therefore, is a potential game (See Section 2.3.1).

This crucial result allows us prove the convergence of behaviour in the population game with boundedly rational agents learning in accordance with Erev and Roth (1998) reinforcement learning (See Proposition 2.2).

The second point is that any 2×2 symmetric pair game that satisfies the anti-coordination condition, Condition 2.1, has a fully interior ESS.

We shall see that these characteristics do not necessarily hold for all symmetric anti-coordination pair games that satisfy Condition 2.1 and are larger than the 2×2 case. Specifically, although any finite population game with random matching consisting of a double symmetric pair game can be shown to admit a potential function, for games with dimension three and greater this precludes the existence of a fully interior mixed strategy ESS in the pair game. Furthermore, even if we relax the double symmetry condition on the pair game, the anti-coordination condition does not guarantee that the mixed strategy ESS of the pair game is indeed fully interior.

S×S Symmetric Pair Games

We can see from our study of the complete class of 2×2 pair games that the interesting case is that of the anti-coordination game. Given this observation there is value in determining if some of the same results apply to finite population games with larger symmetric anti-coordination pair games.

Consider the generic symmetric $S \times S$ anti-coordination pair game, satisfying Condition 2.1, with payoff matrix (2.8).

If this is the case, we can normalise matrix A by subtracting $a_{\alpha\alpha}$ from each column α, for $\alpha = 1, 2, \ldots, S$, without effecting the incentive dynamics, and arrive at the following equivalent matrix:

$$A' = \begin{pmatrix} 0 & (a_{12} - a_{22}) & \cdots & (a_{1S} - a_{SS}) \\ (a_{21} - a_{11}) & 0 & \cdots & (a_{2S} - a_{SS}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{S1} - a_{11}) & (a_{S2} - a_{22}) & \cdots & 0 \end{pmatrix} \tag{2.10}$$

where all payoff entries off the diagonal and strictly greater than zero, i.e. $(a_{\beta\alpha} - a_{\alpha\alpha}) > 0$ for every α and β, $\alpha \neq \beta$.[7] Now let us consider the conditions for which there exists a unique, interior, symmetric mixed strategy equilibrium to the symmetric $S \times S$ anti-coordination pair game. First we must begin with the following definition.

Definition 2.4 *Let Ω denote the $(S-1) \times (S-1)$ dimensional matrix where each element of Ω, denoted by $\omega_{\alpha\beta}$, is a linear combination of elements from A.*

$$\omega_{\alpha\beta} = \left(a_{\alpha\beta} - a_{(\alpha+1)\beta}\right) - \left(a_{\alpha S} - a_{(\alpha+1)S}\right) \tag{2.12}$$

The definition of Ω is relevant because it allows us to explicitly state the conditions on A that ensure the existence of a unique, interior, symmetric mixed strategy equilibrium to the symmetric $S \times S$ anti-coordination pair game. We begin by stating our two conditions on Ω (and therefore on A).

Condition 2.2

$$\det \Omega \neq 0 \tag{2.13}$$

This condition ensures that a mixed strategy equilibrium exists, although it could exist on one of the boundaries. It is clear that in the case of the 2×2 anti-coordination

[7] Note that, as opposed to the generic symmetric 2 x 2 game, the normalised payoff matrix for the generic symmetric 3 x 3 game is not necessarily doubly symmetric. The symmetric 3 x 3 game is doubly symmetric if and only if $a_1 = a_3 - a_{22}$, $a_2 = a_5$ and $a_4 = a_6$, i.e.

$$A' = \begin{pmatrix} 0 & a_1 & a_2 \\ a_1 & 0 & a_4 \\ a_2 & a_4 & 0 \end{pmatrix} \tag{2.11}$$

pair game $\det \Omega \neq 0$; in fact it is always negative. This is not necessarily the case for larger symmetric pair games.

Condition 2.3 *For* $\alpha = 1, 2, \ldots (S-1)$

$$\begin{aligned} &\sum_{\beta=1}^{S-1} C_{\alpha\beta}(\Omega)\left(a_{(\alpha+1)S} - a_{\alpha S}\right) > 0 \quad \text{if } \det \Omega > 0 \\ &\sum_{\beta=1}^{S-1} C_{\alpha\beta}(\Omega)\left(a_{(\alpha+1)S} - a_{\alpha S}\right) > 0 \quad \text{if } \det \Omega < 0 \end{aligned} \tag{2.14}$$

where $C_{\alpha\beta}(\Omega)$ *denotes the* (α, β) *cofactor of* Ω.

This condition ensures that the symmetric mixed strategy is interior and is equivalent to the condition on cofactors of an enlarged matrix as set out in Borkar, Jain, and Rangarajan (2002).

We then follow by stating our proposition and proving that, if indeed these two conditions are met, then the symmetric $S \times S$ anti-coordination pair game has a unique, interior, symmetric mixed strategy equilibrium.

Proposition 2.1 *The symmetric two player* $S \times S$ *anti-coordination pair game, satisfying Condition 2.1 with payoff matrix (2.8), has only one mixed strategy Nash equilibrium which is symmetric if, and only if, Condition 2.2 is satisfied. Furthermore, the unique, symmetric mixed strategy equilibrium is a fully interior symmetric mixed strategy equilibrium if Condition 2.3 is satisfied.*[8]

[8] In addition it can be shown that symmetric mixed strategy equilibrium is the only Nash equilibrium of the symmetric two player $S \times S$ anti-coordination pair game if suitable conditions relating to the relative size of elements of the payoff matrix are satisfied.

Proof We begin by calculating the symmetric mixed strategy equilibrium of the two person game with payoff matrix A. By definition this occurs when the expected payoff from playing each pure strategy is equal.

Therefore, the symmetric mixed strategy Nash equilibrium is precisely the solution to the following simultaneous system:

$$(A\mathbf{x})_1 = (A\mathbf{x})_2 = \cdots = (A\mathbf{x})_S \tag{2.15a}$$

$$x_1 + x_2 + \cdots + x_S = 1 \tag{2.15b}$$

$$x_\alpha > 0 \text{ for } \alpha = 1, 2, \ldots, S \tag{2.15c}$$

As pointed put out by Hofbauer and Sigmund (1998), there exists one or no such solution to this simultaneous system (2.15). Now note that:

$$(A\mathbf{x})_\alpha = \sum_{\gamma=1}^{S} a_{\alpha\gamma} x_\gamma \tag{2.16}$$

represents the expected payoff from playing pure strategy α given $\mathbf{x}$. Since there are only S strategies, we can substitute x_S with $1 - \sum_{\gamma=1}^{S-1} x_\gamma$ and re-express (2.16) in the following terms:

$$(A\mathbf{x})_\alpha = \sum_{\gamma=1}^{S-1} (a_{\alpha\gamma} - a_{\alpha S}) x_\gamma + a_{\alpha S}$$

So the solution to the simultaneous system (2.15) occurs when:

$$\sum_{\gamma=1}^{S-1} (a_{\alpha\gamma} - a_{\alpha S}) x_\gamma + a_{\alpha S} = \sum_{\gamma=1}^{S-1} (a_{\beta\gamma} - a_{\beta S}) x_\gamma + a_{\beta S}$$

for all $\alpha, \beta \in S$, $\alpha \neq \beta$. We can reduce the number of conditions by setting $\beta = \alpha + 1$ for $\alpha = 1, 2, \ldots, S-1$, namely:

$$\sum_{\gamma=1}^{S-1} (a_{\alpha\gamma} - a_{\alpha S}) x_\gamma + a_{\alpha S} = \sum_{\gamma=1}^{S-1} \left(a_{(\alpha+1)\gamma} - a_{(\alpha+1)S}\right) x_\gamma + a_{(\alpha+1)S}$$

Following some simplification we arrive at:

$$\sum_{\gamma=1}^{S-1} \left[\left(a_{\alpha\gamma} - a_{(\alpha+1)\gamma}\right) - \left(a_{\alpha S} - a_{(\alpha+1)S}\right)\right] x_\gamma = a_{(\alpha+1)S} - a_{\alpha S}$$

Now note that for $1 \leq \alpha, \gamma \leq S-1$, $\alpha, \gamma \in \mathbb{N}$, $\left(a_{\alpha\gamma} - a_{(\alpha+1)\gamma}\right) - \left(a_{\alpha S} - a_{(\alpha+1)S}\right)$ is the (α, γ) element of the $(S-1) \times (S-1)$ dimensional matrix Ω. Therefore we have:

$$\left(\Omega \acute{\mathbf{x}}\right)_\alpha = a_{(\alpha+1)S} - a_{\alpha S}$$

If Condition 2.2 is satisfied, then Ω has an inverse and therefore there exists a unique solution for x_α for $\alpha = 1, 2, , S-1$, namely:

$$x_\alpha = \frac{1}{\det \Omega} \sum_{\gamma=1}^{S-1} C_{\alpha\gamma}(\Omega) \left(a_{(\alpha+1)S} - a_{\alpha S}\right) \tag{2.17}$$

If Condition 2.3 holds then:

$$x_\alpha > 0$$

for all $\alpha = 1, 2, , S-1$ and the symmetric pair game with payoff matrix (2.8) has a unique interior mixed strategy Nash equilibrium. ■

2.3 Learning in Finite Population Games

In this section we set out our methodological procedure to study finite population games with random matching and boundedly rational agents who learn in accordance with the Erev and Roth (1998) reinforcement learning model. To achieve this goal, we first need to write down the expected motion of the ith player's strategy adjustment. To achieve this task we begin by defining player i's payoff increment.

Let $\hat{\sigma}^i\left(\delta_t^i, y_t^{-i}\right)$ denote the expected increment to player i's propensity in period t from taking action δ given the aggregate actions taken by the remaining $N-1$ players, denoted by y_t^{-i}, where y_t^{-i} is a vector strategy profile. Note that the updating rule in the Erev and Roth (1998) reinforcement-learning model is a function of realised payoffs. However, the expected motion of the ith player's δ strategy adjustment will be a function of expected payoff increments. This is quantitatively and qualitatively different from realised payoff increments.

To obtain analytical results from the application of the Erev and Roth (1998) reinforcement learning model to finite population games with random matching, we make use of results from the theory of stochastic approximation. In essence we investigate the behaviour of the stochastic learning model by evaluating its expected motion as $t \longrightarrow \infty$. In the case of the Erev and Roth (1998) learning model defined by the choice rule (2.2) and updating rule (2.3), we can write down the expected motion of the ith player's δ strategy adjustment.

Having accomplished this step, we can now state our main results.

Theorem 2.1 *If agents in the finite population game with random matching consisting of the:*

- 2×2 *or,*
- $S \times S$ *column*

anti-coordination pair game satisfying Conditions 2.1 to 2.3, employ the choice rule (2.2) and reinforcement updating rule (2.3), then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching.

Furthermore, if the population size, N, is larger than a threshold level, $\bar{N}$, then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching that corresponds to the mixed strategy equilibrium of the pair game.

Theorem 2.2 *If agents in the finite population game with random matching consisting of the $S \times S$ doubly symmetric anti-coordination pair game satisfying Conditions 2.1 to 2.2, employ the choice rule (2.2) and reinforcement updating rule (2.3), then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching.*

Furthermore, if the population size, N, is larger than a threshold level, $\bar{N}$, then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching that corresponds to the mixed strategy equilibrium of the pair game.[9]

Conjecture 2.1 *If agents in the finite population game with random matching consisting of the anti-coordination pair game satisfying Conditions 2.1 to 2.3, employ the choice rule (2.2) and reinforcement updating rule (2.3), then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching.*

Furthermore, if the population size, N, is larger than a threshold level, $\bar{N}$, then, with probability one, the Erev and Roth (1998) reinforcement learning process converges to a pure Nash equilibrium of the one-shot finite population game with random matching that corresponds to the mixed strategy equilibrium of the pair game.

In order to prove results of this type, we need to study the convergent behaviour of the discrete time stochastic process, describing the expected strategy adjustment of player i's choosing each of the actions available to them. In essence we need to investigate the limit of this process as $t \longrightarrow \infty$.

[9] The statement of Theorem 2.2 differs from Theorem 2.1 because, for doubly symmetric pair games larger than 2×2, the unique symmetric mixed strategy Nash equilibrium of the pair game is not fully interior; that is, it is on one of the boundaries of the strategy space. In these games the symmetric mixed strategy equilibrium involves mixing over some, but not all, of the available pure strategies.

To accomplish this task, we need to establish two main results: one positive and one negative. Drawing these two results together will allow us to confirm our main results, Theorems 2.1 and 2.2, and Conjecture 2.1, outlined above.

The first result we need to establish is the positive convergence result. Here we use a result of Benaïm (1999, Corollary 6.6) to show that the stochastic process describing the expected strategy adjustment of player i's choosing each of the actions will, in the limit as $t \longrightarrow \infty$, converge asymptotically to one of the fixed points of the adjusted replicator dynamics.

The second result we need to establish is the negative convergence result. We employ two results of Hopkins and Posch (2005, Proposition 2 and 3) to demonstrate that the stochastic process describing the expected strategy adjustment of player i's choosing each of the actions will not converge to any fixed points that do not correspond to a Nash equilibria of the one-shot finite population game with random matching or to any corresponding Nash equilibria that are unstable under the adjusted replicator dynamics.

Combined, these two results imply that the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions will converge to the set of strict pure strategy equilibria of the one-shot population game with random matching.

2.3.1 Positive Convergence

In this stage we show that the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions converges with probability one to one of the fixed points of the standard replicator dynamics.

Consider for a moment the behaviour of the following stochastic process (Benveniste, Métivier, and Priouret 1990):

$$x_{t+1} - x_t = \gamma_t f\left(x_t\right) + \gamma_t \eta_t\left(x_t\right) + O\left(\left[\gamma_t\right]^2\right) \tag{2.18}$$

where x_t lies in $\mathbb{R}^N$, $E\left[\eta_t\left(x_t\right)|x_t\right] = 0$ and γ_t defines the nature of the gain in this adaptive process. For our purposes γ_t is interpreted as the step size of the learning algorithm. In our analysis we wish to study the generic convergence properties of stochastic processes of this form as $t \longrightarrow \infty$.

The nature of the gain is important in determining what inferences can be made about the behaviour of (2.18) in the limit. The stronger results from the theory of stochastic approximation apply to adaptive algorithms with decreasing gain; that is stochastic processes with decreasing step size.

Definition 2.5 *The stochastic process (2.18) is said to have decreasing gain if,*

$$\sum_t \left(\gamma_t\right)^{\alpha} < \infty \text{ for some } \alpha > 1 \text{ where } \sum_t \gamma_t = +\infty$$

For example a common step size of $\gamma_t = 1/t$ would ensure that (2.18) has decreasing gain. It emerges that as $t \longrightarrow \infty$, there is a close relationship between the behaviour of stochastic processes (2.18) with a decreasing gain and the mean or averaged ordinary differential equation of the stochastic process.

$$\dot{x} = f(x) \tag{2.19}$$

In particular it can be shown via Benaïm (1999, Corollary 6.6) that if (2.19) meets certain criteria, the stochastic process (2.18) must converge with probability one to one of the fixed points of the mean or averaged ordinary differential equation (2.19).

Theorem 2.3 (Benaïm (1999, Corollary 6.6)) *If the dynamic process (2.19) admits a strict Lyapunov function and processes a finite number of fixed points, then with probability one the stochastic process (2.18) converges to one of these fixed points.*

We now have a method of illustrating that the long-run behaviour of boundedly rational agents, adjusting their strategies according to the Erev and Roth (1998) reinforcement learning model, in the finite population game with random matching converges to one of the fixed points of the mean or averaged differential equation (2.19) associated with the vector of players' expected strategy adjustments.

In order to apply this general result, we must first identify the mean or averaged differential system associated with players' expected strategy adjustment. Furthermore, it must be shown that the mean or averaged differential system admits a strict Lyapunov function. Finally, we must establish that the mean or averaged differential system possesses a finite number of isolated fixed points. In the next three subsections we aim to demonstrate just that.

The Joint Dynamic System

One might hope that the standard replicator dynamics represent the mean or averaged differential system derived from the discrete time stochastic process. Unfortunately, the standard replicator dynamics do not for two simple reasons. First, in the Erev and Roth (1998) reinforcement learning model the step size is endogenous. That is, it is determined by the accumulation of payoffs and thus is not exogenously fixed. Second, the step size is not a scalar.

In order to account for these discrepancies, let us introduce a common step size of $\gamma_t = 1/t$ and N new variables μ_t^i, such that:

$$\mu_t^i = \frac{t}{Q_t^i}$$

We can now substitute $\gamma_t \mu_t^i$ for $1/Q_t^i$ in our discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions and arrive at a *corrected* expected motion of the ith player's strategy adjustment.

Given Condition 2.1, suitable additions to columns can be made to allow all payoffs in the symmetric pair game, and therefore the finite population game, to be positive. This ensures that choice probabilities are well defined, and it follows that μ_t^i is bounded away from zero. Furthermore, since $\mu_t^i = t/Q_t^i$ equals the inverse of the average payoff in the limit as $t \to \infty$, it follows that the associated mean or averaged differential equation (2.19) associated with the corrected discrete time stochastic process is equivalent to the adjusted replicator dynamics in equilibrium. This is extremely useful because there are a variety of results in the literature on the equilibrium behaviour of the adjusted replicator dynamics (see Fudenberg and Levine 1998, Hopkins 2002). We shall revisit some of these findings later in proving Theorems 2.1 and 2.2, and establishing our evidence for Conjecture 2.1.

Because each μ_t^i varies over time, we require a further set of N equations describing the discrete time stochastic process of μ_t^i. Using the method we previously employed to write player i's expected strategy adjustment of each action, we need to find the expected change player i's step size.

Once the mean or averaged differential equation, derived from the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions, has been corrected for the endogenous and non-scalar step size, we arrive at a corrected mean or averaged differential. This system consists of $2N\,(S-1)$ differential equations with $2N\,(S-1)$ endogenous variables describing the evolution of player i's strategy adjustment in the finite population game with random matching. Let us refer to this as the joint dynamic system.

Admission of a Strict Lyapunov Function

We must show that the associated mean or averaged ordinary differential system, the joint dynamic system, admits a strict Lyapunov function. Let us begin with some definitions.

Definition 2.6 *Let (2.19) be an ordinary differential equation defined on some subset Y of $\mathbb{R}^N$. Let $V : Y \to \mathbb{R}$ be a continuously differentiable function. Furthermore, let $\bar{y}$ be a fixed point of $V(y)$. $V(y)$ is a Lyapunov function if,*

$$\dot{V}(y) \geq 0, \qquad \forall \quad y \in Y \textit{ and} \tag{2.20a}$$

$$\dot{V}(\bar{y}) = 0, \qquad \forall \quad \bar{y} \in \theta \tag{2.20b}$$

where θ is the set of fixed points of (2.19).

Definition 2.7 *A strict Lyapunov function is a Lyapunov function $V(y)$ such that:*

$$\dot{V}(y) > 0, \qquad \forall \quad y \notin \theta \tag{2.21}$$

In general it can be difficult and time consuming to identify a suitable Lyapunov function for a particular system. It is often a process of trial and error. An approach to this aspect of the problem developed in the existing literature on the convergence of learning models in games has been to explicitly derive a suitable function for $V(x_1)$ and then show that it admits a strict Lyapunov function. In theory, but not always in practice, this can be accomplished by first assuming that $V(x_1)$ admits a strict Lyapunov function. If this is the case, then the partial derivative $\partial V(x_1)/\partial x_1^i$ represents the expected payoff increment to player i from Strategy 1.

Proposition 2.2 *The finite population game with random matching consisting of 2×2 symmetric pair games admits a strict Lyapunov function, a potential function, and therefore is a potential game.*

Proof The replicator dynamics for a 2×2 population game consisting of the symmetric 2×2 pair game admits the Lyapunov function $V_0(x_1^i)$.

$$V\left(x_1^i\right) = \sum_{i=1}^{N} x_1^i \left\{ (a_1 + a_2) \left[\frac{1}{2} \left(\frac{1}{N-1} \sum_{j \neq i} x^j \right) - \frac{a_2}{(a_1 + a_2)} \right] \right\} \tag{2.22}$$

$$\frac{\partial V\left(x_1^i\right)}{\partial x_1^i} = (a_1 + a_2) \left[\left(\frac{1}{N-1} \sum_{j \neq i} x^j \right) - \frac{a_2}{(a_1 + a_2)} \right] \tag{2.23}$$

Furthermore, the system admits a strict Lyapunov function since it is easily verified that for any fixed point of the replicator dynamics $\dot{V}\left(x_1^i\right) = 0$ and for all other points

$\dot{V}(x_1^i) > 0.$

$$\begin{aligned}
\dot{V}\left(x_1^i\right) &= \frac{\partial V\left(x_1^i\right)}{\partial x_1^i} \cdot \dot{x}_1^i \\
&= (a_1 + a_2)\left[\left(\frac{1}{N-1}\sum_{j\neq i} x^j\right) - \frac{a_2}{(a_1+a_2)}\right] \cdot \dot{x}_1^i \\
&= x_1^i\left(1 - x_1^i\right)\left\{(a_1+a_2)\left[\left(\frac{1}{N-1}\sum_{j\neq i} x^j\right) - \frac{a_2}{(a_1+a_2)}\right]\right\}^2 \geq 0 \quad (2.24)
\end{aligned}$$

■

This is a very important result for finite population games with random matching comprising of 2×2 symmetric pair games. Critically, the result depends on the fact that the payoff matrix of the pair game is symmetric. It is possible to extend this result to all population games with random matching that involve doubly symmetric pair games.

Proposition 2.3 *Any finite population game with random matching comprising a doubly symmetric pair game is a potential game.*

Proof Given that any potential game admits a potential function, the admission of a potential function is equivalent to the admission of a strict Lyapunov function. Consider the following function:

$$\Pi = \sum_{i=1}^{N} \mathbf{x}^i A \bar{\mathbf{x}}^{-i} \qquad (2.25)$$

where $A = A^T$, $\mathbf{x}^i = (x_1^i, x_2^i, ..., x_N^i)$, $\bar{\mathbf{x}}^{-i} = \left(\bar{x}_1^{-i}, \bar{x}_2^{-i}, ..., \bar{x}_N^{-i}\right)$ and $\bar{x}_s^{-i} = \frac{1}{N-1}\sum_{j\neq i} x_s^j$.

If $\Pi(\mathbf{x}^i)$ represents a strict Lyapunov function for the finite population game with random matching, then Conditions (2.20) and (2.21) will be satisfied. Note that:

$$\frac{\partial \Pi}{\partial \mathbf{x}^i} = A\bar{\mathbf{x}}^{-i} + \sum_{j \neq i}^{N} \mathbf{x}^j A \left(\frac{1}{N-1}\right) \tag{2.26}$$

By assumption $A = A^T$, therefore $\mathbf{x}^j A = A\mathbf{x}^j$,

$$\begin{aligned}
\frac{\partial \Pi}{\partial \mathbf{x}^i} &= A\bar{\mathbf{x}}^{-i} + \sum_{j \neq i}^{N} A\mathbf{x}^j \left(\frac{1}{N-1}\right) \\
&= A\bar{\mathbf{x}}^{-i} + A\left(\frac{1}{N-1}\sum_{j \neq i}^{N} \mathbf{x}^j\right) \\
&= A\bar{\mathbf{x}}^{-i} + A\bar{\mathbf{x}}^{-i} \\
&= (2A)\,\bar{\mathbf{x}}^{-i}
\end{aligned}$$

Since $A = A^T$, $\bar{A} = 2A$ is a doubly symmetric matrix:

$$\frac{\partial \Pi}{\partial \mathbf{x}^i} = \bar{\mathbf{x}}^{-i}\bar{A}$$

It is easily verified that $\Pi(\mathbf{x}^i)$ satisfies both Conditions (2.20) and (2.21),

$$\begin{aligned}
\dot{\Pi}(\mathbf{x}^i) &= \frac{\partial \Pi}{\partial \mathbf{x}^i}\dot{\mathbf{x}}^i \\
&= \bar{\mathbf{x}}^{-i}\bar{A}\dot{\mathbf{x}}^i \\
&= \bar{\mathbf{x}}^{-i}\bar{A}\mathbf{x}^i\left[\left(A\bar{\mathbf{x}}^{-i}\right) - \mathbf{x}^i A\bar{\mathbf{x}}^{-i}\right] \\
&= \bar{\mathbf{x}}^{-i}\mathbf{x}^i\left(\bar{A}\cdot A\right)\bar{\mathbf{x}}^{-i} - \bar{\mathbf{x}}^{-i}\mathbf{x}^i\left(\bar{A}\cdot A\right)\mathbf{x}^i\bar{\mathbf{x}}^{-i} \geq 0
\end{aligned}$$

and, therefore, any finite population game with random matching comprising of a doubly symmetric pair game is a potential game. ■

Furthermore, it can be shown that all finite population games with random matching comprising of column anti-coordination pair games, admit a potential function.

Definition 2.8 *A column anti-coordination pair game is defined as a symmetric two-player pair game satisfying Condition 2.1 where all elements of each column off the diagonal are equal.*

$$A = \begin{pmatrix} a_{11} & a_2 & \cdots & a_S \\ a_1 & a_{22} & \cdots & a_S \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_{SS} \end{pmatrix} \tag{2.27}$$

Proposition 2.4 *Any finite population game with random matching comprising a column anti-coordination pair game is a potential game.*

Proof Begin by noting that the payoff matrix for any column anti-coordination pair game (2.27) can re-scaled by subtracting the diagonal element in each column from that column, and then subtracting the common value off the diagonal in each column for every column. We arrive at a the following diagonal matrix:

$$\begin{pmatrix} a_{11} - a_1 & 0 & \cdots & 0 \\ 0 & a_{22} - a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{SS} - a_S \end{pmatrix}$$

By assumption the payoff matrix for the column anti-coordination pair game (2.27) satisfies Condition 2.1. Therefore, each element on the diagonal is less than zero. The structure of the payoff matrix now defines a pure congestion game, and it follows directly from the result of Monderer and Shapley (1996, Theorem 3.1) that the column anti-coordination pair game is a potential game and, therefore, admits a potential function. Furthermore, any finite population game with random matching comprising a column anti-coordination pair game is a potential game. ■

We do not have such a result for the larger class of symmetric pair games that are not doubly symmetric or column anti-coordination pair games. We do, however, have the following conjecture:

Conjecture 2.2 *All finite population games with random matching comprising of anti-coordination pair games satisfying Condition 2.1 are isomorphic to congestion games.*

The ability to prove Conjecture 2.2 to be true would be a significant step towards providing a general result for the application of the Erev and Roth (1998) reinforcement learning model to finite population games with random matching and symmetric anti-coordination pair games. This follows directly from the result of Monderer and Shapley (1996, Theorem 3.1).

Fixed Points of the Joint Dynamic System

Definition 2.9 *The fixed points are the rest points of the joint dynamic system.*

We now need to show that the joint dynamic system possesses a finite number of isolated fixed points. The fixed points of the $N(S-1)$ equations describing the evolution of the step size occur when either the step size is zero or a value which is an inverse function of the average play of the remaining $(N-1)$ players and expected payoff increments. By assumption, all payoffs are positive, therefore $\bar{\mu}^i$ is bounded away from zero. This means that the fixed points of the joint dynamic system with $\bar{\mu}^i = 0$ are always unstable (see Hopkins 2002, Duffy and Hopkins 2005) and, therefore, are never asymptotic outcomes. We can now concentrate on the case where the step size is an inverse function of the average play of the remaining $(N-1)$ players and expected payoff increments.

Consider the first $N(S-1)$ equations of the joint dynamic system. Once we substitute for $\bar{\mu}^i$ and multiply both sides by the denominator, we arrive at a reduced form of the joint dynamic system consisting of $N(S-1)$. This implies that the fixed points of the joint dynamic system are exactly the same as those under the adjusted replicator dynamics and, consequently, the standard replicator dynamics. The characterisation of the fixed point of the standard replicator dynamics is well known (see Weibull 1995) and consists of the union of all pure states and Nash equilibria of the underlying game.

We also need to demonstrate that the set of all Nash equilibria of the underlying game, the one-shot finite population game with random matching, is finite. Note that there are three types of Nash equilibria of the underlying game, namely:

- **Pure Strategy Nash Equilibria**

  Nash equilibria where all players play a pure strategy.

- **Symmetric Mixed Strategy Nash Equilibria**

  Nash equilibria where all players play a mixed strategy.

- **Asymmetric Mixed Strategy Nash Equilibria**

  Nash equilibria where some players play a pure strategy and the remaining play a mixed strategy.

If the set of all Nash equilibria of the underlying game is finite, the joint dynamic system will possess a finite number of fixed points. To be absolutely clear, the fixed points of the joint dynamic system consist of the following:

- **Pure strategy Nash equilibria**

  These are the pure states of the joint dynamic system that correspond to the pure strategy Nash equilibria of the underlying game.

- **Symmetric mixed strategy Nash equilibrium**

  This is the full interior state of the joint dynamic system that corresponds to the symmetric mixed strategy Nash equilibria of the underlying game. That is, the Nash equilibrium where all players play a mixed strategy best response.

- **Asymmetric mixed strategy Nash equilibria**

  These are boundary states of the joint dynamic system that correspond to asymmetric mixed strategy Nash equilibria of the underlying game. By boundary states we mean those where a subset of the N players play a mixed strategy best response, while the remainder play a pure strategy.

- **Fixed points that are not Nash equilibria**

  Not all fixed points of the joint dynamic system correspond to Nash equilibria of the underlying game. There are pure states of the joint dynamic system that do not correspond to pure strategy Nash equilibria of the underlying game. Note that it is not possible to have interior fixed points or fixed points on some boundary of the state space of the joint dynamic system that do not correspond to Nash equilibria of the underlying game.

Positive Convergence

Once we have shown that the joint dynamic system admits a strict Lyapunov function and that it possesses a finite number of fixed points which are identical to those of the standard replicator dynamics, we can apply Theorem 2.3, Benaïm (1999, Corollary 6.6). Application of this theorem proves that the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions converges to one of the fixed points of the standard replicator dynamics.

2.3.2 Negative Convergence

In the second part of the proof of the main results, we show that the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions does not converge to any equilibria corresponding to Nash equilibria of the underlying game which are unstable under the adjusted replicator dynamics or equilibria that do not correspond to a Nash of the underlying game. We show this in two steps.

First, we show that the stability properties of a fixed point of the joint dynamic system are entirely determined by the stability properties of the corresponding fixed point under the adjusted replicator dynamics. We then determine the stability properties of the Nash equilibria under the adjusted replicator dynamics. We conclude that only the strict pure strategy Nash equilibria are stable under the adjusted replicator dynamics. We then employ Hopkins and Posch (2005, Proposition 2) to show that the discrete time stochastic process, which describes the expected strategy adjustment of player i's choosing each action, cannot converge to any fixed point that is unstable under the adjusted replicator dynamics.

Second, we employ Hopkins and Posch (2005, Proposition 3) to demonstrate that the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions cannot converge to any fixed point that does not correspond to a Nash equilibria of under the underlying game. Therefore, we have our negative convergence result.

Unstable Equilibria in the Adjusted Replicator Dynamics

Definition 2.10 *A fixed point of the joint dynamic system is unstable if its linearisation evaluated at* $\bar{x}$ *has at least one eigenvalue with a positive real part.*

Theorem 2.4 (Hopkins and Posch (2005, Proposition 2)) *Let* $\bar{x}$ *be a Nash equilibrium that is linearly unstable under the adjusted replicator dynamics. Then the Erev and Roth (1998) reinforcement learning process defined by choice rule (2.2) and* *the updating rule (2.3) asymptotically converges to one of these points with probability zero.*

Lemma 2.1 *The stability properties of the fixed points of the joint dynamic system are entirely determined by the stability properties of the corresponding fixed points of the adjusted replicator dynamics.*

What this series of result implies is that in order to determine the stability properties of the fixed points of the joint dynamic system, we only need to study are the stability properties of the fixed points of the adjusted replicator dynamics that correspond to Nash equilibria of the underlying game.

Lemma 2.2 *The fixed points of the adjusted replicator dynamics corresponding to the strict pure strategy Nash equilibria of the one-shot finite population game are asymptotically stable.*

Proof Given that the pure strategy Nash equilibria are strict, they constitute evolutionary stable strategies of the one-shot population game with random matching. By Weibull (1995), all evolutionary stable strategies are asymptotically stable under the replicator dynamics. ■

Lemma 2.2 is very useful as we can show that finite population games with random matching, comprising a symmetric anti-coordination game satisfying certain

Conditions 2.1 to 2.3 plus an additional condition on the structure of the payoff matrix (2.8), will generically admit a strict pure strategy. If there exists a division of the population that is close enough to the mixed strategy Nash equilibrium of the anti-coordination pair game, then there exists a strict pure strategy Nash equilibria in the one-shot finite population game with random matching.

The following theorem demonstrates that the 2×2 finite population game with random matching consisting of the 2×2 anti-coordination pair game admits a strict pure strategy Nash equilibrium in pure strategies.

Theorem 2.5 *Consider the symmetric 2×2 anti-coordination pair game with the following payoff matrix:*

$$A = \begin{pmatrix} 0 & a_1 \\ a_2 & 0 \end{pmatrix}, \text{ where } a_j > 0 \qquad \forall j \tag{2.28}$$

There exists a strict pure strategy Nash equilibrium in the finite population game with N players, $N \in 2\mathbb{N}$, who are randomly matched to play the pair game if there exists $\phi \in \mathbb{N}$ such that

$$x_1^*(N-1) < \phi < x_1^*(N-1) + 1 \tag{2.29}$$

where $(x_1^*, (1 - x_1^*))$ denotes the symmetric mixed strategy Nash equilibrium of the pair game.

Proof Begin by assuming that all players in a population of N players are playing pure strategies. Let $\boldsymbol{\varphi}=(\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2)$ be the proportion of players playing pure strategy 1 and 2 respectively. Given that there are only two pure strategies, $\boldsymbol{\varphi}_2 = (1-\boldsymbol{\varphi}_1)$. Note that $N\boldsymbol{\varphi}_1, N\boldsymbol{\varphi}_2 \in \mathbb{N}$ represents the number of players playing pure strategy 1 and 2 respectively in the finite population game.

In order to demonstrate that there exists a pure strategy Nash equilibrium in the finite population game with N players, $N \in 2\mathbb{N}$, who are randomly matched to play the symmetric 2×2 anti-coordination pair game, it is necessary to demonstrate the existence of a pure strategy profile in the population such that each player playing each pure strategy has no incentive to deviate.

We begin by stating the expected payoff to players playing each strategy and then the payoff they would expect to receive if they deviated to the other pure strategy.

The expected payoff to a player playing pure strategy 1 is equal to the probability that they will be matched with a player playing each pure strategy times the payoff they would receive from each match.

Formally, the expected payoff to a player playing pure strategy 1 given the pure strategies played by the remaining $(N-1)$ players is[10]

$$E\left[\pi_1|\boldsymbol{\varphi}\right] = \left(\frac{N\left(1-\boldsymbol{\varphi}_1\right)}{N-1}\right)a_1$$

[10]Note that the probability of being matched with a player playing each pure strategy is a function of $(N-1)$ because the player in question knows which one of the two pure strategies they are playing themselves.

Similarly, the expected payoff to pure strategy 1 players deviating to pure strategy 2 is

$$E\left[\pi_{1\to 2}|\boldsymbol{\varphi}\right] = \left(\frac{N\varphi_1 - 1}{N-1}\right) a_2$$

Therefore, a player playing pure strategy 1 has no incentive to deviate to pure strategy 2 if and only if $E\left[\pi_1|\boldsymbol{\varphi}\right] > E\left[\pi_{1\to 2}|\boldsymbol{\varphi}\right]$, i.e.

$$(a_1 + a_2)\,\varphi_1 < a_1\left(1 - \frac{1}{N}\right) + (a_1 + a_2)\left(\frac{1}{N}\right) \tag{2.30}$$

Now let us consider the incentives for a player playing pure strategy 2. The expected payoff to a player playing pure strategy 2 is

$$E\left[\pi_2|\boldsymbol{\varphi}\right] = \left(\frac{N\varphi_1}{N-1}\right) a_2$$

Similarly, the expected payoff to players playing pure strategy 2 deviating to pure strategy 1 is

$$E\left[\pi_{2\to 1}|\boldsymbol{\varphi}\right] = \left(\frac{N\,(1-\varphi_1) - 1}{N-1}\right) a_1$$

Therefore, a player playing pure strategy 2 has no incentive to deviate to pure strategy 1 if and only if $E\left[\pi_2|\boldsymbol{\varphi}\right] > E\left[\pi_{2\to 1}|\boldsymbol{\varphi}\right]$, i.e.

$$(a_1 + a_2)\,\varphi_1 > a_1\left(1 - \frac{1}{N}\right) \tag{2.31}$$

Since by assumption, $a_1 + a_2 > 0$, we can divide both inequality (2.30) and (2.31) by $a_1 + a_2$, which implies that for a pure strategy Nash equilibrium to be supported in the finite population game

$$\frac{a_1}{a_1+a_2}\left(1-\frac{1}{N}\right) < \varphi_1 < \frac{a_1}{a_1+a_2}\left(1-\frac{1}{N}\right) + \left(\frac{1}{N}\right)$$

Given that $x_1^* = \frac{a_1}{a_1+a_2}$, it follows that:

$$x_1^*\left(1-\frac{1}{N}\right) < \varphi_1 < x_1^*\left(1-\frac{1}{N}\right) + \left(\frac{1}{N}\right)$$

In conclusion a pure strategy Nash equilibrium in the finite population game may be supported if there is no incentive for players playing either pure strategy to deviate from playing their chosen pure strategy. In other words a strict pure strategy Nash equilibrium exists in the population game if there exists a division of the population of players playing each pure strategy that is sufficiently close to the symmetric mixed strategy Nash equilibrium of the pair game, i.e. if there exists $\phi \in \mathbb{N}$, such that

$$x_1^*(N-1) < \phi < x_1^*(N-1)+1, \text{ where } 0 \leq \phi \leq N \tag{2.32}$$

We have proven the existence of a strict pure strategy Nash equilibrium in the finite population game consisting of N players, $N \in 2\mathbb{N}$, who are randomly matched to play the pair game with payoff matrix (2.28). ■

Theorem 2.5 says that if there exists a division of the population that is close enough to the mixed strategy Nash equilibrium $(x_1^*, 1-x_1^*)$ of the symmetric 2×2 anti-coordination pair game, then there exists a strict pure strategy Nash equilibrium in the finite population game.[11]

[11]Note that if this is not the case, i.e. in a non-generic set-up where $x_1^*(N-1)$ and $x_1^*(N-1)+1 \in \mathbb{N}$, then there exists an infinite number of asymmetric mixed strategy Nash equilibria where all players except one play an appropriate pure strategy and the final player is indifferent between each pure strategy. In this case any pure strategy Nash equilibria are non-strict.

We have proven that the one-shot 2×2 finite population game with random matching comprising of the 2×2 anti-coordination pair game admits strict Nash equilibria in pure strategies. Given this observation it would be interesting to determine if this result extends to finite population games with random matching comprising of larger symmetric anti-coordination pair games.

The next stage of our analysis is to state the set of conditions which will ensure that the one-shot finite population game with random matching composed of the $S \times S$ anti-coordination pair game admits a strict pure strategy Nash equilibrium.

Theorem 2.6 *Let* $\mathbf{n} = (n_1, n_2, n_3, ..., n_S)$*, where* $n_\alpha \in \mathbb{N}$ *and* $\sum^S n_\alpha = N$*, denote a pure strategy profile of the finite population game consisting of* N *players randomly matched into pairs to play the symmetric* $S \times S$ *anti-coordination pair game with the following payoff matrix:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{pmatrix}$$

The pure strategy profile $\mathbf{n} = (n_1, n_2, n_3, ..., n_S)$ *is a strict Nash equilibrium of the one shot* $S \times S$ *anti-coordination finite population game if, for every pure strategy* α*, such that* $n_\alpha \neq 0$*, the following* $\frac{S(S-1)}{2}$ *double-inequality conditions are satisfied for all*

$\beta \neq \alpha$:

$$a_{\alpha\alpha} - a_{\beta\alpha} < (A\mathbf{n})_\alpha - (A\mathbf{n})_\beta < a_{\alpha\beta} - a_{\beta\beta} \tag{2.33}$$

Proof By assumption $\mathbf{n} = (n_1, n_2, n_3, ..., n_S)$, where $n_\alpha \in \mathbb{N}$ and $\sum^S n_\alpha = N$, denotes a pure strategy profile of the finite population game consisting of N players randomly matched into pairs to play the symmetric two player $S \times S$ anti-coordination pair game with payoff matrix A.

Given the pure strategies played by the other players in the population, the payoff a player playing pure strategy α would expect to receive, denoted $E[\pi_\alpha | \mathbf{n}]$, is equal to the payoff a player playing pure strategy α would receive if they were matched with a player playing pure strategy γ times the probability that a player playing pure strategy α is matched with a player playing pure strategy γ summed over all possible matches. Therefore,

$$\begin{aligned} E[\pi_\alpha | \mathbf{n}] &= \sum_{\gamma=1}^{S} a_{\alpha\gamma} \Pr(\alpha \text{ player is matched with a } \gamma \text{ player} | \mathbf{n}) \\ &= a_{\alpha 1}\left(\frac{n_1}{N-1}\right) + \cdots + a_{\alpha(\alpha-1)}\left(\frac{n_{\alpha-1}}{N-1}\right) \\ &\quad + a_{\alpha\alpha}\left(\frac{n_\alpha - 1}{N-1}\right) \\ &\quad + a_{\alpha(\alpha+1)}\left(\frac{n_{\alpha+1}}{N-1}\right) + \cdots + a_{\alpha S}\left(\frac{n_S}{N-1}\right) \\ &= \sum_{\gamma \neq \alpha}^{S} \frac{a_{\alpha\gamma} n_\gamma}{N-1} + \frac{a_{\alpha\alpha}(n_\alpha - 1)}{N-1} = \sum_{\gamma=1}^{S} \frac{a_{\alpha\gamma} n_\gamma}{N-1} - \frac{a_{\alpha\alpha}}{N-1} \end{aligned}$$

Therefore,

$$E[\pi_\alpha | \mathbf{n}] = \frac{(A\mathbf{n})_\alpha - a_{\alpha\alpha}}{N-1}$$

Assuming the remaining players in the population continue to play their respective pure strategies, the payoff a player playing pure strategy α deviating to pure strategy β would expect to receive, denoted $E\left[\pi_{\alpha\to\beta}\left|\mathbf{n}\right.\right]$, is

$$\begin{aligned}
E\left[\pi_{\alpha\to\beta}\left|\mathbf{n}\right.\right] &= \sum_{\gamma=1}^{S} a_{\beta\gamma}\Pr\left(\beta \text{ player is matched with a } \gamma \text{ player}\left|\mathbf{n}\right.\right) \\
&= a_{\beta 1}\left(\frac{n_1}{N-1}\right) + \cdots + a_{\beta(\alpha-1)}\left(\frac{n_{\alpha-1}}{N-1}\right) \\
&\quad + a_{\beta\alpha}\left(\frac{n_\alpha - 1}{N-1}\right) \\
&\quad + a_{\beta(\alpha+1)}\left(\frac{n_{\alpha+1}}{N-1}\right) + \cdots + a_{\beta S}\left(\frac{n_S}{N-1}\right) \\
&= \sum_{\gamma\neq\alpha}^{S}\frac{a_{\beta\gamma}n_\gamma}{N-1} + \frac{a_{\beta\alpha}\left(n_\alpha - 1\right)}{N-1} = \sum_{\gamma=1}^{S}\frac{a_{\beta\gamma}n_\gamma}{N-1} - \frac{a_{\beta\alpha}}{N-1}
\end{aligned}$$

Therefore,

$$E\left[\pi_{\alpha\to\beta}\left|\mathbf{n}\right.\right] = \frac{\left(A\mathbf{n}\right)_\beta - a_{\beta\alpha}}{N-1}$$

Thus, a player playing strategy α has no incentive to deviate to strategy β if and only if

$$E\left[\pi_\alpha\left|\mathbf{n}\right.\right] > E\left[\pi_{\alpha\to\beta}\left|\mathbf{n}\right.\right]$$

It follows that the pure strategy profile $\mathbf{n}$ is a strict pure strategy Nash equilibrium if, for every pure strategy α, the following conditions are satisfied for all $\beta \neq \alpha$:

$$\left(A\mathbf{n}\right)_\alpha - \left(A\mathbf{n}\right)_\beta > \left(a_{\alpha\alpha} - a_{\beta\alpha}\right)$$

Given that the term $\left(A\mathbf{n}\right)_\alpha - \left(A\mathbf{n}\right)_\beta$ appears in the condition that requires that a player playing strategy α has no incentive to deviate to strategy β and vice versa, we can express the $S(S-1)$ inequality conditions as the $\frac{S(S-1)}{2}$ double inequality conditions (2.33). ■

We have now stated the set of double inequality conditions that must be satisfied to ensure the existence of a strict pure strategy Nash equilibrium in the population game.[12]

Condition 2.4 *For any payoff matrix A that satisfies Condition 2.1, and for all α, β such that $|\alpha - \beta| > 1$, let*

$$a_{\alpha\beta} \leq \begin{cases} \sum_{\gamma=\alpha}^{\beta-1} a_{(\gamma+1)\gamma} & \text{if } \alpha > \beta \\ \sum_{\gamma=\alpha}^{\beta-1} a_{\gamma(\gamma+1)} & \text{if } \alpha < \beta \end{cases}$$

This condition ensures that for any symmetric two player $S \times S$ anti-coordination pair games the sum of the elements on the upper or lower diagonal is always greater to or equal to the elements in the upper or lower corners of the payoff matrix A, i.e. those elements of payoff matrix A not in either diagonal adjacent to the centre diagonal. This condition is important because it allows us to generalise our proof of the existence of strict pure strategy Nash equilibria in $S \times S$ anti-coordination finite population games for pair game payoff matrices with dimensions higher than 3×3.

[12]For doubly symmetric pair games not satisfying Condition 2.3, it is actually a subset of the $S(S-1)$ inequality conditions that must be satisfied to support a strict Nash equilibria in the corresponding finite population game with random matching.

At this point it is interesting to note that although Borkar, Jain, and Rangarajan (2002) provide evidence that larger than 3×3 anti-coordination finite population games admit strict pure strategy Nash equilibria, they never explicitly state that Condition 2.4 must be satisfied for these $S \times S$ anti-coordination pair games. The following theorem will show that this is a necessary condition in order to reduce the $\frac{S(S-1)}{2}$ double inequality conditions (2.33) to the $S-1$ double inequality conditions for $\alpha = 1, 2, ..., S-1$.

We now state our next result which allows us to reduce the $\frac{S(S-1)}{2}$ double inequality conditions (2.33) to the $S-1$ double inequality conditions for $\alpha = 1, 2, ..., S-1$.

Theorem 2.7 *If for payoff matrix A, Condition 2.4, is satisfied we can represent the $\frac{S(S-1)}{2}$ double inequality conditions (2.33) with the following $S-1$ double inequality conditions for $\alpha = 1, 2, ..., S-1$:*

$$a_{\alpha\alpha} - a_{(\alpha+1)\alpha} < (A\mathbf{n})_{\alpha} - (A\mathbf{n})_{(\alpha+1)} < a_{\alpha(\alpha+1)} - a_{(\alpha+1)(\alpha+1)} \qquad (2.34)$$

Proof Begin by letting $\beta = \alpha+1$, where $\alpha = 1, 2, ..., S-1$, in the double inequality conditions (2.33). This gives us the $S-1$ double inequality conditions (2.34), where $\alpha = 1, 2, ..., S-1$.

It should be apparent that this smaller set of double inequality conditions do not, on their own, ensure that the original set of double inequality conditions (2.33) hold for every pure strategy α, such that $n_\alpha \neq 0$, and every pure strategy β, such that $\beta \neq \alpha$ and $\beta \neq \alpha + 1$. In fact the $S-1$ double inequality conditions (2.34) only ensure that the double inequality conditions (2.33) hold for all α and β where $\beta = \alpha + 1$.

Let us assume that this is not the case and that $\beta = \alpha + \delta$, where $\delta \in \mathbb{N}$. To support a strict pure strategy Nash equilibrium, the following condition, among others, would have to be satisfied:

$$a_{\alpha\alpha} - a_{(\alpha+\delta)\alpha} < (A\mathbf{n})_\alpha - (A\mathbf{n})_{(\alpha+\delta)} < a_{\alpha(\alpha+\delta)} - a_{(\alpha+\delta)(\alpha+\delta)} \tag{2.35}$$

By summing over the $S-1$ double inequality conditions (2.34) we can re-write (2.35) as

$$\sum_{\gamma=\alpha}^{\beta-1} \left(a_{\gamma\gamma} - a_{(\gamma+1)\gamma}\right) < \sum_{\gamma=\alpha}^{\beta-1} \left((A\mathbf{n})_\gamma - (A\mathbf{n})_{(\gamma+1)}\right) < \sum_{\gamma=\alpha}^{\beta-1} \left(a_{\gamma(\gamma+1)} - a_{(\gamma+1)(\gamma+1)}\right) \tag{2.36}$$

From the double inequality conditions (2.33) we know that

$$\begin{aligned} a_{\alpha\alpha} - a_{\beta\alpha} &\geq \sum_{\gamma=\alpha}^{\beta-1} \left(a_{\gamma\gamma} - a_{(\gamma+1)\gamma}\right) \\ a_{\alpha\beta} - a_{\beta\beta} &\leq \sum_{\gamma=\alpha}^{\beta-1} \left(a_{\gamma(\gamma+1)} - a_{(\gamma+1)(\gamma+1)}\right) \end{aligned} \tag{2.37}$$

Given that (2.37) are satisfied by Condition 2.4 for any payoff matrix which satisfies Condition 2.1, it follow that the $\frac{S(S-1)}{2}$ double inequalities must hold (2.35) if the $S-1$ double inequality conditions (2.34) are satisfied. ■

Since not all n_γ are independent, we can rewrite the above double inequality conditions (2.34) in terms of $\acute{\mathbf{n}} = (n_1, n_2, n_3, ..., n_{S-1})$ by substituting for $n_s =$

$N-(n_1+n_2+n_3+\cdots+n_{S-1})$. This gives us the following $S-1$ double inequality conditions:

$$\begin{aligned}&\left(a_{\alpha\alpha}-a_{(\alpha+1)\alpha}\right)-\left(a_{\alpha S}-a_{(\alpha+1)S}\right)N\\&<\sum_{\gamma=1}^{S-1}\left(\left(a_{\alpha\gamma}-a_{(\alpha+1)\gamma}\right)-\left(a_{\alpha S}-a_{(\alpha+1)S}\right)\right)n_\gamma<\\&\left(a_{\alpha(\alpha+1)}-a_{(\alpha+1)(\alpha+1)}\right)-\left(a_{\alpha S}-a_{(\alpha+1)S}\right)N\end{aligned}\tag{2.38}$$

Note by Definition 2.4,

$$\Omega\acute{\mathbf{n}}=\sum_{\gamma=1}^{S-1}\left(\left(a_{\alpha\gamma}-a_{(\alpha+1)\gamma}\right)-\left(a_{\alpha S}-a_{(\alpha+1)S}\right)\right)n_\gamma\tag{2.39}$$

where Ω is a $S-1$ dimension square matrix. If Ω has an inverse, we can isolate for $\acute{\mathbf{n}}$ and obtain $S-1$ double inequality conditions bounding n_α. Note that Ω^{-1} existence depends only on the payoff matrix A, and not on N. Furthermore, by Proposition 2.1, Ω does in fact have an inverse since Condition 2.2 must be satisfied. We can now solve explicitly for n_α for $\alpha=1,2,,S-1$.

$$\begin{aligned}&\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}(\Omega)\left(\left(a_{\gamma S}-a_{(\gamma+1)S}\right)N-\left(a_{\gamma(\gamma+1)}-a_{(\gamma+1)(\gamma+1)}\right)\right)\\&<n_\alpha<\\&\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}(\Omega)\left(\left(a_{\gamma S}-a_{(\gamma+1)S}\right)N-\left(a_{\gamma\gamma}-a_{(\gamma+1)\gamma}\right)\right)\end{aligned}$$

$$\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma S}-a_{(\gamma+1)S}\right)N-\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma(\gamma+1)}-a_{(\gamma+1)(\gamma+1)}\right)$$

$$<n_{\alpha}<$$

$$\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma S}-a_{(\gamma+1)S}\right)N-\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma\gamma}-a_{(\gamma+1)\gamma}\right)$$

We can now substitute equation (2.17) from the proof of Proposition 2.1 to arrive at

$$\begin{aligned}
&x_{\alpha}N-\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma(\gamma+1)}-a_{(\gamma+1)(\gamma+1)}\right)\\
&<n_{\alpha}<\\
&x_{\alpha}N-\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(a_{\gamma\gamma}-a_{(\gamma+1)\gamma}\right)
\end{aligned}\tag{2.40}$$

We can then rewrite (2.40) so that we have n_{α} bounded in a finite interval, whose location is dependent on the number of players, N, and whose size is depends only on the payoff matrix A, and not N.

$$x_{\alpha}\left(N-1\right)+\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(\left(a_{\gamma S}-a_{(\gamma+1)S}\right)-\left(a_{\gamma(\gamma+1)}-a_{(\gamma+1)(\gamma+1)}\right)\right)$$

$$<n_{\alpha}<$$

$$x_{\alpha}\left(N-1\right)+\frac{1}{\det\Omega}\sum_{\gamma=1}^{S-1}C_{\alpha\gamma}\left(\Omega\right)\left(\left(a_{\gamma S}-a_{(\gamma+1)S}\right)-\left(a_{\gamma\gamma}-a_{(\gamma+1)\gamma}\right)\right)$$

At this stage we have demonstrated that there are a set of conditions on the payoff matrix of the pair game that guarantee the existence of strict Nash equilibria in the finite population games with random matching.

Finally, we must show that all the remaining Nash equilibria in the one-shot finite population game with random matching are asymptotically unstable under the adjusted replicator dynamics. This set consists of the fully symmetric mixed and asymmetric mixed strategy Nash equilibria in the one-shot population game with random matching and any other pure strategy equilibria of the one-shot finite population game with random matching that do not correspond to the mixed strategy equilibrium of the pair game in frequencies.

First, we show that all fully symmetric mixed and asymmetric mixed strategy Nash equilibria in the one-shot population game with random matching are asymptotically unstable under the adjusted replicator dynamics. Fortunately, it is not difficult to show that this is the case (see Weibull 1995).

Second, we need to demonstrate that if N is large enough, $N > \bar{N}$, then any other pure strategy equilibria of the one-shot finite population game (that do not correspond to the mixed strategy equilibrium of the pair game in frequencies) are unstable under the adjusted replicator dynamics. Borkar, Jain, and Rangarajan (1998) have made some progress with regard to this issue for 3×3 anti-coordination pair games.

Example 2.4 *This is what the general bounds on n_α for the 3×3 anti-coordination matrix look like:*

$$\begin{aligned} & x_1^*(N-1) + \frac{(a_{13} - a_{12} - a_{23})(a_{23} + a_{32})}{\det \Omega} \\ & < n_1 < \\ & x_1^*(N-1) + \frac{(a_{12} + a_{21})(a_{23} + a_{32})}{\det \Omega} n_1 \end{aligned} \tag{2.41}$$

$$x_2^*(N-1) + \frac{(a_{13}-a_{12}-a_{23})(a_{21}-a_{23}-a_{31})}{\det\Omega}$$
$$< n_2 < \tag{2.42}$$
$$x_2^*(N-1) + \frac{(a_{31}-a_{21}-a_{32})(a_{23}-a_{13}-a_{21})}{\det\Omega}$$

Remark It is useful to mention here that the existence of the two double inequality conditions for the finite population game comprising of the 3×3 anti-coordination game does not necessarily imply a unique, strict, pure strategy Nash equilibrium profile. It would be possible to add further restrictions onto the structure of the payoff matrix on the pair game to ensure the profile is unique.

Non-Nash Fixed Points of the Joint Dynamic System

Theorem 2.8 (Hopkins and Posch (2005, Proposition 3)) *Let $\bar{x}$ be a fixed point of the replicator dynamics which is not a Nash equilibrium. The Erev and Roth (1998) reinforcement learning process, defined by the choice rule (2.2) and the updating rule (2.3), asymptotically converges to one of these points with probability zero.*

Therefore, the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions cannot converge to any fixed point not corresponding to a Nash equilibrium under the underlying game.

Negative Convergence Result

Proposition 2.5 *The discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions converges with probability zero to equilibria corresponding to Nash equilibria of the underlying game which are unstable under the adjusted replicator dynamics or equilibria not corresponding to Nash equilibria of the underlying game.*

Proof The result follows from Theorem 2.4, Hopkins and Posch (2005, Proposition 2), and Theorem 2.8, Hopkins and Posch (2005, Proposition 3). ■

2.4 Simulations

It has not been possible to arrive at a definitive result proving that long-run behaviour of boundedly rational agents, learning in accordance with the Erev and Roth (1998) reinforcement learning model, converges to a pure strategy state in finite population games with random matching comprising of anti-coordination pair games. Therefore, we must now explore further, through the use of simulations, the convergence properties of the discrete time stochastic process describing the expected strategy adjustment of player i's choosing each of the actions. Borkar, Jain, and Rangarajan (1998) take this approach and have produced evidence from simulations that suggests that in finite population games with random matching, learning does converge to the set of pure strategy equilibria.

Under certain conditions our evidence from simulating the learning environment suggests individual agents tend to specialise. Each pure strategy is played by at least one agent, therefore the population as a whole retains its diversity so as the aggregate outcome in the population mimics the mixed strategy equilibrium of the pair game in frequencies.

In the following three subsections we present some indicative data from our simulations for 2×2 and 3×3 anti-coordination finite population games with random matching.

2.4.1 2×2 Anti-Coordination Finite Population Games

Consider the Hawk-Dove pair game with $c = 6$ and $v = 2$.

		Player 2	
		Fight	*Yield*
Player 1	*Fight*	$-2, -2$	$2, 0$
	Yield	$0, 2$	$1, 1$

Once we normalise the payoff matrix to ensure that all elements are positive we have

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

We know that the symmetric mixed strategy equilibrium in the pair game is $\left(\frac{1}{3}, \frac{2}{3}\right)$. Therefore, with a finite population of six players we would expect two players in the population to converge to the pure strategy 'fight' and four to 'yield'. We can see that in the following simulation with a step size of 0.15 and one-hundred repetitions this is indeed the outcome (See Figure 2.1).

Simulations were run for a whole series of 2×2 anti-coordination finite population games with random matching, and the outcome in all case was that long-run behaviour converged to pure strategy states.

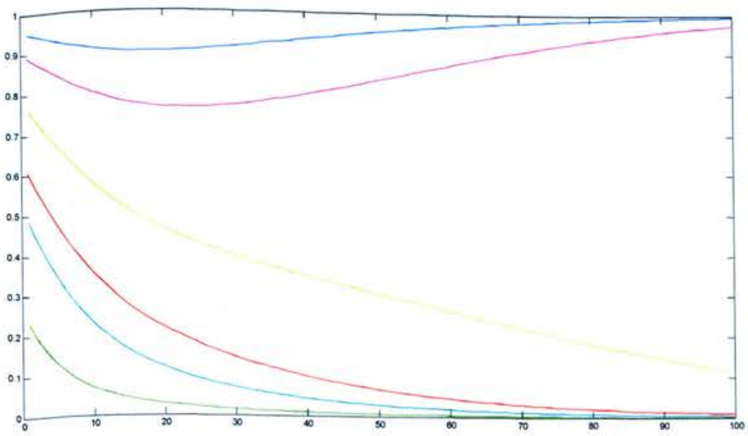

Figure 2.1: Simulation of the Hawk-Dove Finite Population Game

2.4.2 3×3 Anti-Coordination Finite Population Games

Consider the 3×3 anti-coordination pair game with the following normalised payoff matrix:

$$A = \begin{pmatrix} 0 & 108 & 324 \\ 432 & 0 & 108 \\ 108 & 405 & 0 \end{pmatrix}$$

Engle-Warnick and Hopkins (2006) have show in their experimental paper investigating learning with fixed matching that the pair game has a unique interior mixed strategy equilibrium with relative probabilities $(17, 20, 24)/61 \approx (0.279, 0.328.0.393)$. It can be shown that there exists a pure strategy profile in the finite population game with random matching and eighteen players. This profile sees five players converge to playing the first pure action, six to the second and seven to the third; which is what we see in the simulations. An example run with a step size of 0.01 and one-hundred and fifty repetitions can be seen in Figure 2.2.

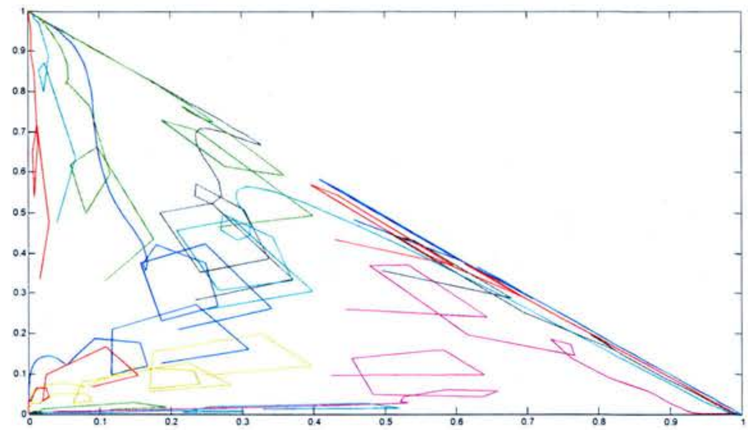

Figure 2.2: Simulation of the 3×3 Anti-Coordination Finite Population Game

As can be seen from these example simulations and other simulations that we have run, long-run behaviour consistently converges to the pure strategy profile of the one-shot finite population game with random matching.x_2

2.5 Conclusion

The intuition behind our results is that the choice of matching protocol and population setting can transform a strategic game with interior evolutionary stable strategies to one with strict pure strategies which have been shown in the literature to be evolutionarily stable.

The main conjecture implies sorting and is crucially dependent on the fact that the finite population games with random matching that we study are potential games. Although we have demonstrated that 2×2 anti-coordination finite population games are potential games, further work is required to extend these findings to general anti-coordination finite population games. Given the evidence provided by our simulations, we believe that there is a real possibility that a result could be proved which extends to general anti-coordination finite population games.

We have shown that in 2×2 anti-coordination finite population games, learning theory does predict sorting. We believe that this is an important conclusion, which contributes to the literature. We feel that this prediction should be a consideration when analysing data from laboratory experiments in economics on populations games.

Appendix 2.A MATLAB Code

2.A.1 Code for 2×2 Finite Population Game Simulation

```
T=[Number of Periods];
N=[Population Size];
s=[Step Size];

A=[Pair Game Payoff Matrix];

a11=A(1,1);
a12=A(1,2);
a21=A(2,1);
a22=A(2,2);

b1=a11-a21-a12+a22;
b2=a12-a22;

x=zeros(T,N);
xib=zeros(1,N);

x(1,:)=rand(1,N);

for t=1:T
    for i=1:N
```

```
    xib(i)=(sum(x(t,:))-x(t,i))/(N-1);

    end

    x(t+1,:)=x(t,:)+s*x(t,:).*(1-x(t,:)).*(b1*xib+b2);

end

plot(x(1:T,:))
```

2.A.2 Code for 3×3 Finite Population Game Simulation

```
T=[Number of Periods];

N=[Population Size];

s=[Step Size];

A=[Pair Game Payoff Matrix];

a11=A(1,1);

a12=A(1,2);

a13=A(1,3);

a21=A(2,1);

a22=A(2,2);

a23=A(2,3);

a31=A(3,1);

a32=A(3,2);

a33=A(3,3);

b1=(a11-a13)-(a31-a33);

b2=(a12-a13)-(a32-a33);

b3=a13-a33;

b4=(a21-a23)-(a31-a33);

b5=(a22-a23)-(a32-a33);
```

```
b6=a23-a33;

x=zeros(T,N);

y=zeros(T,N);

xib=zeros(1,N);

yib=zeros(1,N);

x(1,:)=rand(1,N)/2;

y(1,:)=rand(1,N)/2;

for t=1:T

    for i=1:N

    xib(i)=(sum(x(t,:))-x(t,i))/(N-1);

    yib(i)=(sum(y(t,:))-y(t,i))/(N-1);

    end

    x(t+1,:)=x(t,:)+s*(x(t,:).*(1-x(t,:)).

        *((b1.*xib+b2.*yib+b3)))-s*(x(t,:).*y(t,:).*((b4.*xib+b5.*yib+b6)));

    y(t+1,:)=y(t,:)+s*(y(t,:).*(1-y(t,:)).

        *((b4.*xib+b5.*yib+b6)))-s*(x(t,:).*y(t,:).*((b1.*xib+b2.*yib+b3)));

end

plot(x(1:T,:),y(1:T,:))
```

Chapter 3

Sorting in Anti-Coordination Finite Population Games: A Laboratory Experiment in Economics

3.1 Introduction

This paper is an experimental study of individual learning in two types of finite population games with random matching under two different information treatments. The two games consist of a Hawk-Dove anti-coordination finite population game and a 3×3 anti-coordination finite population game. Both of these games use random matching and involve pair games that admit a unique symmetric mixed strategy equilibrium. The pair games represent the two player games that are played once subjects are matched into pairs.

Theories of learning are frequently being subjected to tests using data from controlled laboratory experiments with paid human subjects. In our work we contribute to this growing body of literature by testing how well the long-run predictions of the Erev and Roth (1998) model of reinforcement learning, applied to anti-coordination finite populations games with random matching, track the actual behaviour of participants in laboratory environments.

We show that the theory of learning predicts that, under certain conditions, play in finite population games with random matching should not only converge to Nash equilibria, but that it should only converge to pure strategy Nash equilibria of finite population games with random matching.[1] These equilibria correspond, in frequencies, to the symmetric mixed strategy Nash equilibrium of the pair game if

[1]Specifically, we show that this is the case for all finite population games consisting of symmetric 2×2, doubly symmetric $S \times S$ and column $S \times S$ anti-coordination pair games.

the pair game is an anti-coordination game and if the payoff matrix satisfies certain conditions, as discussed in Chapter 2 and to be revisited in Section 3.2.

In the experiments subjects were allocated randomly and anomalously into groups, or finite populations, to play a repeated finite population game with random pairwise matching. In each round of the finite population game, subjects within each group were randomly matched into pairs to play the symmetric two-player anti-coordination game. This process was repeated over one-hundred rounds.

In total we ran six finite population games with random matching consisting of one-hundred rounds for both the 2×2 Hawk-Dove anti-coordination game and the 3×3 anti-coordination game. These twelve games were run under the aggregate information setting, where minimal feedback information was available to subjects. We then ran five finite population games with random matching consisting of one-hundred rounds for both the 2×2 Hawk-Dove anti-coordination game and the 3×3 anti-coordination game under a full information treatment, where more detailed feedback information was available to subjects.

As one would expect, we found that behaviour varied across the games. In aggregate in the populations, average frequencies of play converged very quickly to the evolutionary stable strategies in the corresponding pair game. That is, average frequencies of play looked very similar to those generated by Nash equilibrium play. Previous experiments have reported similar findings; see for example, Friedman

(1996), Cheung and Friedman (1998). Up to now, however, no studies that we are aware of have explicitly looked for evidence that repeated play leads to convergence to pure strategy Nash equilibrium in finite population games with random matching.

We investigate the hypothesis that, given sufficient repeated play and adequate feedback, individual participants in experimental finite population games should learn equilibrium behaviour. As Duffy and Hopkins (2005) point out, this type of claim naturally raises the following two questions: what in practice is a sufficient number of repetitions to allow participants to learn equilibrium behaviour and what is adequate information to allow this process to occur?

With regards to the first question, in our experimental design we have chosen to repeat the interactions one hundred times. This number of repetitions was chosen based on the advice of other experimenters in the field, and because it is the most rounds of play that could be carried out in a session without unduly pressurising subjects to make choices, while still ensuring that we had a common number of rounds across sessions and treatments. This number of rounds also allowed us to complete our sessions within the time limit detailed by the recruitment rules of the laboratory where we conducted our experiments.

Concerning the second question about adequate information, we followed the experimental design as set out in Duffy and Hopkins (2005). In their research they show that a simple reinforcement learning model and a more sophisticated model, referred to as stochastic fictitious play or hypothetical reinforcement learning have the same long-run predictions, given a suitable number of repetitions. The two models differ in terms of their sophistication, but under both play must converge to a pure strategy equilibrium, commonly referred to as a sorting outcome. A sorting outcome occurs when some players always choose one of the alternatives and others always chose other alternatives, where no players play a mix over several of the alterative actions. The reader should be reminded that even if these learning models accurately predict decision behaviour, there is no guarantee that we will be able to see these predicted outcomes in the time available in our experiment. It is entirely possible that these predicted outcome would only be observable after several hundred, if not several thousand, repetitions.

With this in mind, we adopted the following information treatments. First, an aggregate information treatment where subjects were given information on the number and share of members of their group who chose each available action. Second, a full information treatment where subjects were given the precise decisions made by each member of their group without identifying any group member individually. In Duffy and Hopkins's (2005) investigations of market entry games, evidence suggests that convergence occurs faster under the full information treatment than the aggregate information treatment.

Previous experimental investigations of finite population games with random matching have focused on testing whether the mixed strategy equilibrium of the pair game characterises the average frequencies of play in finite population games. The data from some of these experiments (Friedman 1996) suggests a more heterogeneous outcome, with some participants mixing between several choices and some playing a pure strategy. However, the average frequencies of play in the finite population games do converge to the mixed strategy equilibrium of the pair game.

In another relevant experiment in economics, Erev and Rapoport (1998) reported that, in market entry games, for which learning theory also predicts convergence to a sorting outcome, the speed of convergence towards Nash equilibrium levels increases the more feedback information is available.

To summarise, in both the 2×2 Hawk-Dove finite population game with random matching and the 3×3 anti-coordination finite population game with random matching, and under both information treatments, we are testing the hypothesis that individuals' play will converge to pure strategies. In other words we expect to see a sorting outcome in the finite population. Furthermore, we are testing the hypothesis that if more feedback information is available to individual participants, then convergence to the predicted outcome will be faster.

Overall, evidence has been found to support our first hypothesis in the 2×2 aggregate information treatment. However, the strength of this evidence diminishes as we move to the 3×3 aggregate information treatment and both the 2×2 and 3×3

games under the full information treatment. We do, however, still see some evidence in all four treatments of players converging to playing pure strategies. We hypothesise that the complexity of the games themselves might be limiting our ability to observe convergence to sorting outcomes in the time available.

The next section, Section 3.2, of the paper details our theoretical predictions of learning theory in finite population games with random matching consisting of 2×2 and 3×3 anti-coordination pair games. In Section 3.3 we detail the experimental design and procedure. In Section 3.4 we present our results. In the final section, Section 3.5, we discuss our results in further detail and provide some concluding remarks.

3.2 Theoretical Predictions

Before we outline our theoretical predictions for the long-run behaviour of individuals in anti-coordination finite population games with random matching who learn in accordance with the Erev and Roth (1998) model of reinforcement learning, let us formally define the finite population game and anti-coordination finite population game.

Definition 3.1 *The finite population game with random matching is defined as the infinitely repeated game where, in each period, all members of a finite population of players are randomly matched into pairs to play a symmetric two-player simultaneous move game. All players realise their payoffs and are returned to the population pool, to play the same symmetric two-player game again in the next period against another randomly drawn player.*

Definition 3.2 *The anti-coordination finite population game with random matching is defined as the finite population game with random matching where the pair game satisfies the following conditions:*

Condition 3.1 *The payoff matrix of the pair game is diagonally sub-dominant. That is, for payoff matrix A,*

$$a_{\alpha\alpha} < a_{\alpha\beta} \text{ for all } \alpha \neq \beta \text{ and } \alpha, \beta \in S \tag{3.1}$$

where α and β denote pure strategies from the set of all pure strategies, S.

Condition 3.2 *The pair game admits an interior mixed strategy Nash equilibrium.*

Condition 3.3 *The elements of the payoff matrix of the pair game satisfy the following inequalities for all α, β where $|\alpha - \beta| > 1$,*

$$a_{\alpha\beta} \leq \begin{cases} \sum_{\gamma=\alpha}^{\beta-1} a_{(\gamma+1)\gamma} & \text{if } \alpha > \beta \\ \sum_{\gamma=\alpha}^{\beta-1} a_{\gamma(\gamma+1)} & \text{if } \alpha < \beta \end{cases} \tag{3.2}$$

where $a_{\alpha\beta}$ denotes the payoff to a row if that player chose pure strategy α and the column player chose pure strategy β.

For 2×2 symmetric pair games satisfying Condition 3.1, Conditions 3.2 and 3.3 are automatically satisfied. For larger symmetric pair games satisfying Condition 3.1, this is not necessarily the case, and Conditions 3.2 and 3.3 must be checked independently.

Recall from our study in Chapter 2 that the long-run behaviour of boundedly rational agents in finite population games with random matching, who learn in accordance with the Erev and Roth (1998) model of reinforcement learning, converges to the set of pure strategy Nash profiles of the one-shot finite population game with random matching if,

- the symmetric pair game satisfies Conditions 3.1 to 3.3; and,
- the one-shot finite population game with random matching admits a potential function and is, therefore, a potential game.

In Chapter 2 we proved that any finite population game with random matching consisting of a 2×2 symmetric pair game admits a potential function. Furthermore, we showed that any finite population game with random matching consisting of a doubly symmetric pair game admits a potential function. However, any doubly symmetric pair game larger than 2×2 does not satisfy Condition 3.2. This has important consequences for our theoretical predictions in these experiments as it implies that we will only have definitive predictions for the 2×2 treatment cases.

Our convergence result for the long-run behaviour of boundedly rational agents learning in finite population games with random matching, in accordance with the Erev and Roth (1998) model of reinforcement learning, is crucially dependent on our ability to show that the finite population game with random matching admits a Lyapunov function. By definition all potential games admit a potential function and, given the continuous set of mixed strategies in finite population games with random matching, we can express the potential function as a smooth function with respect to the strategy space. From here it is easy to prove that the joint dynamic system describing the expected strategy adjustment of player i's choosing each of the actions in finite population games with random matching admits a Lyapunov function (see Chapter 1, Lemma 1.2).

While we cannot apply this theoretical prediction with the same rigor to the 3×3 treatment cases, there is still value in testing these game treatments in an experimental setting. Recall the simulation results presented in Chapter 2, which provided evidence that models of reinforcement learning do lead to sorting outcomes in 3×3 anti-coordination finite population games with random matching.

Finally, despite the fact that the learning model considered here predicts no effect from the availability of additional information, experimental evidence from Duffy and Hopkins's (2005) indicates that the provision of additional information leads to a higher tendency towards sorting in market entry games. In our research we wish to investigate the validity of this observation in alternative game treatments. This brings us to our two hypotheses:

Hypothesis 3.1 In anti-coordination finite population games with random matching, the long-run behaviour of boundedly rational agents learning in accordance with a model of reinforcement learning, converges to the set of strict pure strategy equilibria. That is, reinforcement learning predicts sorting.

Hypothesis 3.2 In anti-coordination finite population games with random matching, the speed of convergence to the set of pure strategy equilibria increases with the availability of information. That is, convergence to a sorting outcome happens faster in the full information treatments than the aggregate information treatments.

3.3 Experimental Design and Procedure

3.3.1 Experimental Design

The experiment is a 2×2 design consisting of a game treatment (2×2 and 3×3 anti-coordination finite population game) and an information treatment (aggregate or full information). As discussed in the previous section, both the 2×2 and 3×3 anti-coordination pair games that constitute the respective finite population games with random matching have unique, interior, symmetric mixed strategy equilibria. Furthermore, both the pair games satisfy Conditions 3.1 to 3.3. Given the structure of the payoff matrices of the pair games, it can be proven that both the 2×2 and 3×3 anti-coordination finite population games admit a strict pure strategy Nash profile, which approximates the unique, interior, symmetric mixed strategy equilibria of the pair game.

	A	B
A	80	260
B	360	120

Figure 3.1: Payoff Matrix for the 2×2 Anti-Coordination Pair Game

Consider the payoff matrix for the 2×2 (Hawk-Dove) anti-coordination pair game (see Figure 3.1). There are three Nash equilibria for the 2×2 anti-coordination pair game, two asymmetric pure strategy Nash equilibria and one symmetric mixed strategy Nash equilibrium with relative probabilities $(1, 2)/3 \approx (0.333, 0.667)$. In the finite population game with six players, there is a unique, interior, symmetric mixed strategy Nash equilibrium where each player in the population plays the same mixed strategy Nash equilibrium as in the pair game. There are also a number of asymmetric mixed strategy Nash equilibria where a subset of the population play pure strategies and the remainder play mixed.

Both the symmetric and asymmetric mixed strategy equilibria in the finite population game with random matching are unstable under the adjusted replicator dynamics. In addition there is a set of strict pure strategy Nash profiles where two members of the population play action **A** and four play action **B.** Since the elements of this set of pure strategy profiles are strict Nash, they are asymptotically stable under the adjusted replicator dynamics. Furthermore, given that the 2×2 anti-coordination pair game with random matching is a potential game, the long-run behaviour of subjects learning according to the Erev and Roth (1998) model of reinforcement learning, converges to the set of strict pure strategy Nash profiles of the 2×2 anti-coordination pair game with random matching; that is, where two members of the population play action **A** and four play action **B.**

	A	B	C
A	80	200	260
B	480	120	180
C	160	280	100

Figure 3.2: Payoff Matrix for the 3×3 Anti-Coordination Pair Game

Consider the payoff matrix for the 3×3 anti-coordination pair game (see Figure 3.2). There is a unique, interior, symmetric mixed strategy Nash equilibrium in the 3×3 anti-coordination pair game with relative probabilities $(2, 3, 1)/6 \approx (0.333, 0.5, 0.167)$. In fact this is the only Nash equilibrium in this particular pair game. In the finite population game with six players, there is a unique, interior, symmetric mixed strategy Nash equilibrium where each player in the population plays the same mixed strategy Nash equilibrium as in the pair game. There are also a number of asymmetric mixed strategy Nash equilibria where a subset of the population play pure strategies and the remainder play mixed.

Both the symmetric and asymmetric mixed strategy equilibria in the finite population game with random matching are unstable under the adjusted replicator dynamics. In addition there is a set of strict pure strategy Nash profiles where two members of the population play action **A**, three play action **B** and one plays action **C.** Since the elements of this set of pure strategy profiles are strict Nash, they are asymptotically stable under the adjusted replicator dynamics. Furthermore, our conjecture that the 3×3 anti-coordination pair game with random matching is a congestion game and, therefore, a potential game, means that the long-run behaviour of subjects learning in accordance with the Erev and Roth (1998) model of reinforcement learning, converges to the set of strict pure strategy Nash profiles of the 3×3 anti-coordination pair game with random matching; that is, where two members of the population play action **A** and three play action **B** and one plays action **C.**

3.3.2 Experimental Procedure

We ran the four different treatments over eight sessions, two sessions for each treatment. The four treatments consisted of a 2×2 and a 3×3 aggregate information treatment and a 2×2 and a 3×3 full information treatment. Thirty-six subjects (six groups of six) participated in each of the 2×2 and 3×3 aggregate information treatments, and thirty subjects (five groups of six) participated in each of the 2×2 and 3×3 full information treatments. No subject participated in more than one session.

The rules of the game were common knowledge to all participants (see Sections 3.B.1 to 3.B.3 in Appendix 3.B and Sections 3.C.1 to 3.C.4 in Appendix 3.C for the full experimental script and treatment instructions) who were given complete information regarding their opponent's behaviour in each round of play. Depending on the information treatment being tested, participants were also given information regarding the behaviour of all members of their group, including themselves.

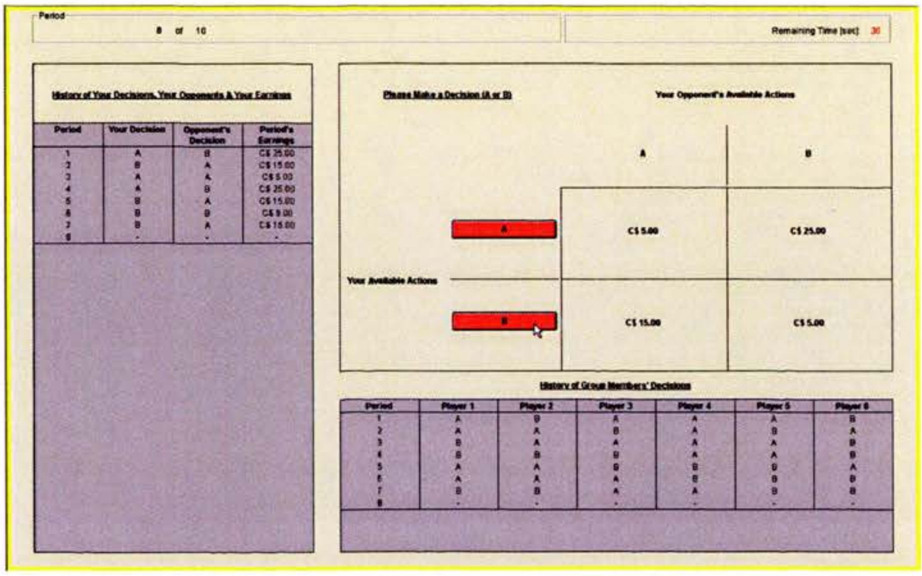

Figure 3.3: An example 2×2 full information screen layout

In the upper-right hand of their computer screen, the subjects were presented with either a 2×2 or 3×3 payoff matrix depending on the game treatment. Participants were then asked to make a decision and click the relevant button to record their decision. The computer interface presented payoff information to all participants as if they were the row player, with the screen revealing only the subject's payoff in each cell; opponents payoffs were not revealed.[2] On the left side of the computer

[2]Note that most theories of learning, including those to be considered here, do not assume any knowledge of opponents' payoff.

screen, subjects could scroll through the entire history of their decisions, as well as the decisions of the opponent to whom they were matched within each round of the experiment. At the bottom right of the screen, subjects were shown, depending on treatment type, either aggregate information or full information on the play of all the members of their group in each round.

In the aggregate information treatment, for both the 2×2 and 3×3 game treatments, a tally and share figure was presented, detailing the number of group members who chose each available action in each round. In the full information treatments, a full decision history for each member of each group was presented. This information was presented in a format that ensured that no individual participant could be individually identified. Again, participants could scroll through the entire history of their group's decisions. The historical information was then updated after each round of decision making.

Subjects were told that they would make one-hundred decisions over the course of one-hundred rounds, and that other participants with whom they were randomly grouped together, and would subsequently be matched with in each round, would remain the same throughout the entire session. They were also told that they would be paid for ten periods of play to be randomly determined by the computer. The ten randomly selected periods would be drawn independently for each participant.

Participants were not quizzed to ascertain if they understood how to read the

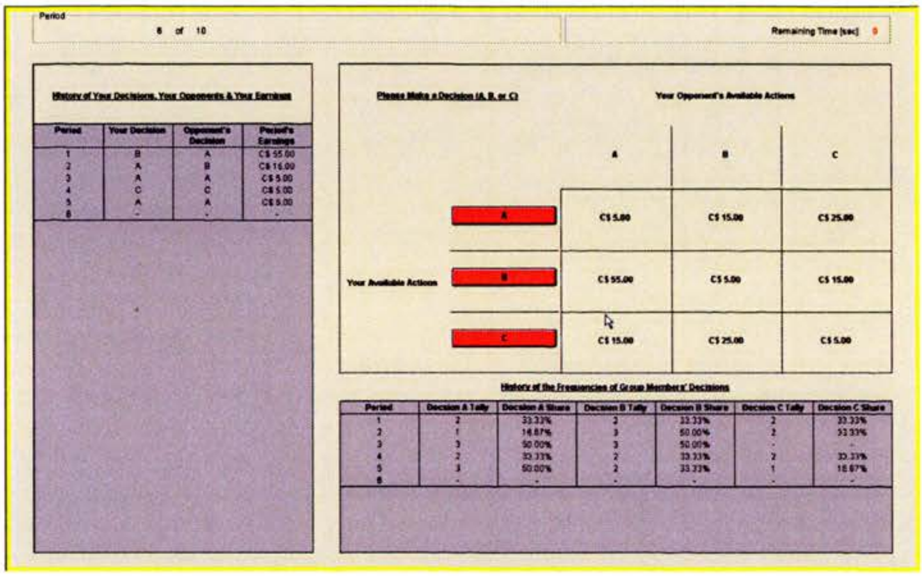

Figure 3.4: An example 3×3 aggregate information screen layout

earnings table prior to the sessions commencing. This decision was taken for three reasons. First, the subject pool was a mature one, with many of the participants having had previous experience as subjects in experiments in economics. Second, we were investigating individual behaviour in a symmetric game, therefore, each participant in each game treatment faced the same earnings table. Finally, it was felt that ample instruction was provided in explaining how to read the payoff matrix on a whiteboard prior to the start of each session.[3]

A total of one-hundred and thirty-two subjects, all of whom were English-speaking university students in Montreal, participated in the four experimental treatments.

[3]Furthermore, the payoff matrix remained on the whiteboard throughout the session so participants could refer to it as and when required.

Twenty-two groups of six subjects, of whom forty-eight percent were male and fifty-two percent were female, were recruited using the ORSEE Recruitment System (Greiner 2004). The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007). The experiments were run in August and September 2007 at the Bell Experimental Laboratory for Commerce and Economics at the Centre for Research and Analysis on Organizations (CIRANO). Subjects earned CAD \$10.00 (GBP £5.00) for showing up on time. This show-up fee reflected the fact that the laboratory was not on campus and, therefore, an additional incentive was required to recruit subjects for the sessions.

Participants also earned an average of CAD \$21.04 (ranging between \$13.20 and \$28.80 with a median of \$21.00 and a mode of \$20.20) for the results of their decisions and the decisions of their opponents. Alternative opportunities for work in Montreal pay approximately CAD \$8.00 per hour. Our sessions never lasted more than two hours.

3.4 Experimental Results

For each of the finite population game treatments we present the results in graphical form, see Appendix 3.A. Four figures have been included for each of the twenty-two game treatments.

In the first figure, for all game treatments, we present one or two smaller graphs depicting the number of subjects in each anti-coordination game choosing each action in each period, the red lines. For the 2×2 game treatments, this is achieved using a single graph detailing the number of subjects choosing action **B**. For the 3×3 game treatments, two graphs are presented: one detailing the number of subjects choosing action **A** and the other action **B**. Overlaid on these graphs is the average number of subjects in each game choosing each action in each period, the green lines, and the predicted number of subjects who should choose each action in each period, the blue lines.

In the second figure, six smaller graphs depict the actual decisions subjects made during the course of the session for all game treatments. For the avoidance of confusion, players are denoted by colour. Starting in the first row and first column of the array of six graphs, moving right along the top row and then from left to right along the bottom row, the players in each game are called BLUE, BROWN, GREEN, ORANGE, TURQUOISE and RED.

The third and fourth figures represent different quantities depending on the game treatment. For the 2×2 game treatments, they depict the average play every ten periods on the third figure, and the moving average of play every ten periods on the fourth figure. The colours of the lines correspond to the names of the players. In other words the average play every ten periods and the moving average of play every ten periods of the BLUE player is represented by the blue lines.

For the 3×3 game treatments, it is not meaningful to calculate the average play and the moving average of play every ten periods. Instead, we calculate a scaled distance measure to determine the "purity" of the average play every ten periods and the "purity" of the moving average of play every ten periods for each subject.

$$\text{Purity} = 1 - \left(\frac{(\sqrt[3]{\bar{x}_{\mathbf{A}}} \cdot \sqrt[2]{\bar{x}_{\mathbf{B}}} \cdot \sqrt[6]{\bar{x}_{\mathbf{C}}})}{\frac{1}{36} \cdot 3^{\frac{2}{3}} \cdot 6^{\frac{5}{6}} \cdot \sqrt{2}} \right) \tag{3.3}$$

Note that whenever a player is playing a pure strategy, the "purity" measure equals one. On the other hand, if a player is playing the symmetric mixed strategy, the "purity" measure equals zero.[4]

In the following four subsections (Sections 3.4.1 to 3.4.4), we note the number of subjects who appear to have converged to a pure strategy and to which pure strategy they appear to have converged. In addition we summarise the treatment data by

[4]We employ a similar, appropriately scaled, "purity" function for the 2×2 game treatments in our treatment summary representations.

presenting graphs showing the mean decision purity averaged over 10 fixed rounds for each game under each treatment.

3.4.1 2×2 Aggregate Information Treatment

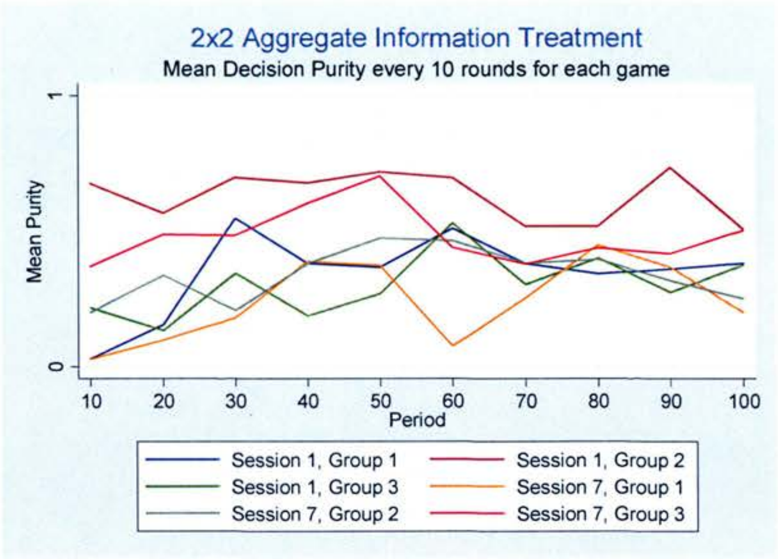

Figure 3.5: 2 × 2 Aggregate Information Treatment: Mean Decision Purity every 10 rounds for each game.

In Session 1, Group 1, we see two players converge to pure strategies, one to **A** and one to **B.** In Session 1, Group 2, we see three players converge to pure strategies, one to **A** and two to **B**. In Session 1, Group 3, we see two players converge to pure strategies, both to **B**. In Session 7, Group 1, we see two players converge to pure strategies, both to **B**. In Session 7, Group 2, we see only one player converge to a pure strategy, **B**. In Session 7, Group 3, we see two players converge to pure strategies, both to **B**.

In the six 2×2 finite population games with aggregate information, we see at least one player in each game converges to pure strategy **B**.

In order to further establish evidence of convergence to pure strategies we now study the average purity across the treatment. In Figure 3.6, each data point represents that mean decision purity every 10 period across the whole treatment.

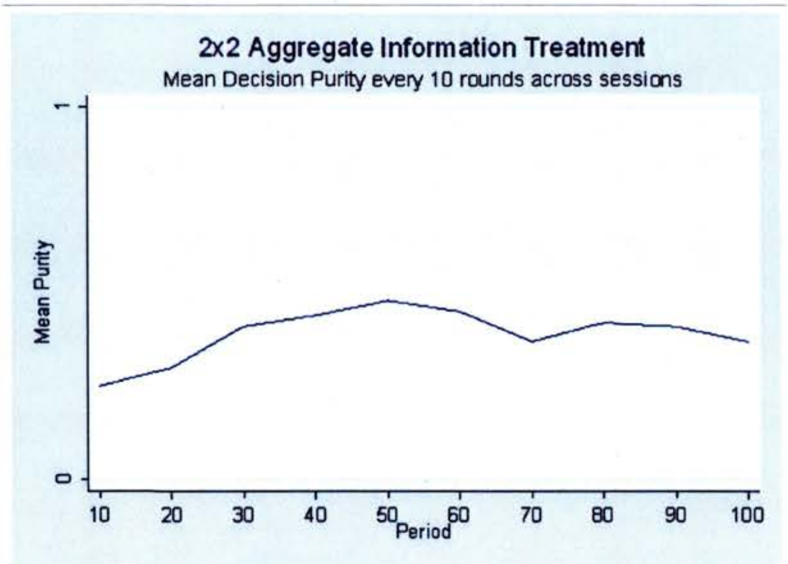

Figure 3.6: Average Putity in 2×2 Aggregate Information Treatment

In the 2×2 aggregate information treatment sessions the mean decision purity increases over time with a coefficient of the linear regression model that is not quite significant at the 10% level of significance. There is little evidence against the null hypothesis that the coefficient is less than or equal to zero (H_o : $Coef$ <= 0 : $p - value = 0.10526426$).

We also investigate the variance of mean decision purity every 10 rounds in each game (see Figure 3.7). We find strong evidence, at the 1% level of significance, that the variance of mean decision purity is decreasing over time. There is strong very strong evidence against the null hypothesis that coefficient is greater than or equal to zero (H_o : $coef >= 0$: $p - value = 0.0080075$). Together, these two pieces of evidence provide encouraging support for the prediction that individual play should converge to pure strategies.

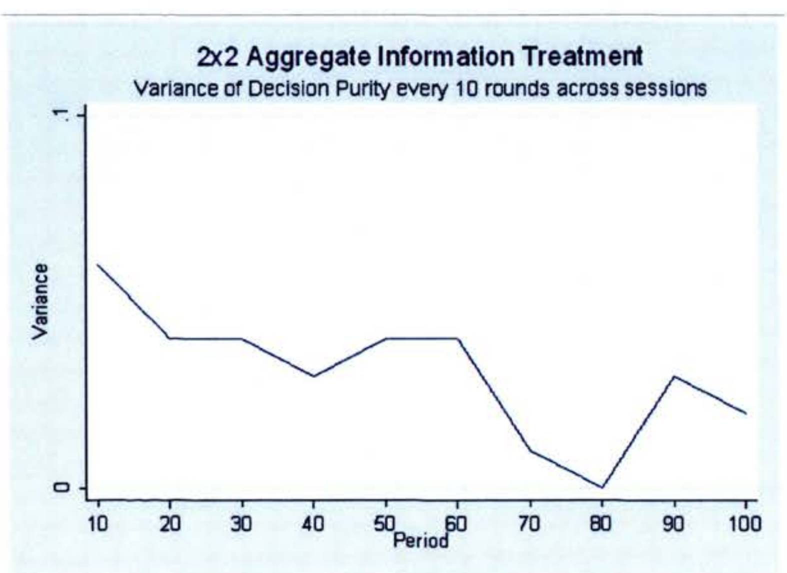

Figure 3.7: Variance of Average Putity in 2 × 2 Aggregate Information Treatment

3.4.2 2×2 Full Information Treatment

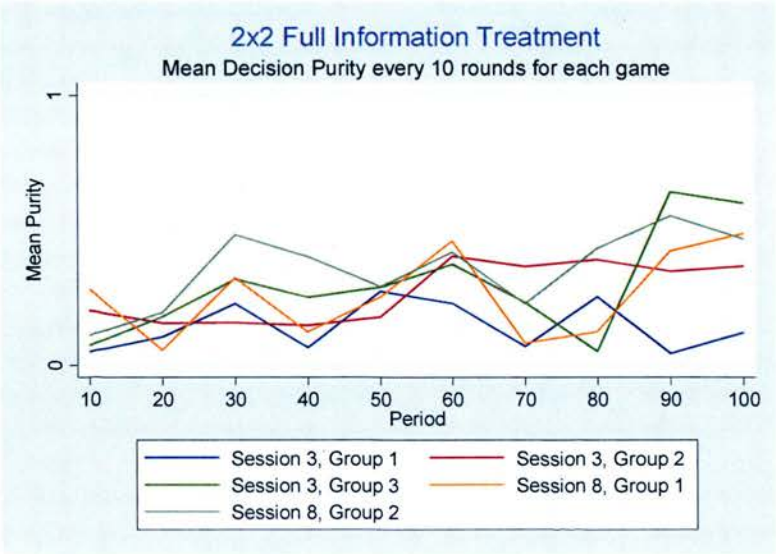

Figure 3.8: 2×2 Full Information Treatment: Mean Decision Purity every 10 rounds for each game.

In Session 3, Group 1, we see only one player converge to a pure strategy, **B.** In Session 3, Group 2, we see two players converge to pure strategies, one to **A** and one to **B**. In Session 3, Group 3, we see only one player converge to a pure strategy, **B**. In Session 8, Group 1, we see only one player converge to a pure strategy, **B**. In Session 8, Group 2, we see only one player converge to a pure strategy, **B**.

In the five 2×2 finite population games with full information, we see at least one player in each game converges to pure strategy **B**.

In order to further establish evidence of convergence to pure strategies we now study the average purity across the treatment. In Figure 3.9, each data point represents that mean decision purity every 10 period across the whole treatment.

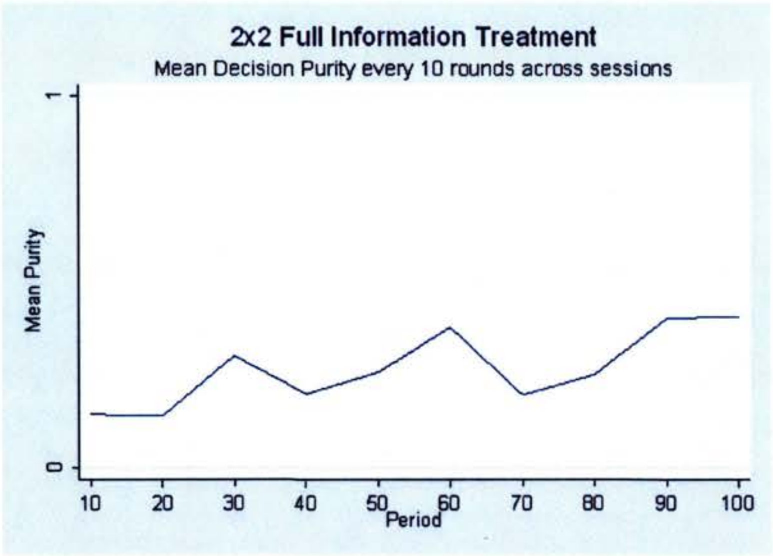

Figure 3.9: Average Putity in 2×2 Full Information Treatment

In the 2×2 full information treatment sessions the mean decision purity increase over time with a coefficient of the linear regressing model that is significant at the 1% level of significance. We can reject the null hypothesis that the coefficient is less than or equal to zero ($H_o : Coef <= 0 : p - value = 0.00778238$).

We also investigate the variance of mean decision purity every 10 rounds in each game (see Figure 3.10). We find no real evidence against the null hypothesis that the variance of mean decision purity is increasing over time ($H_o : Coef >= 0 : p- value =$

0.99120975). Therefore, both the mean decision purity across treatment sessions and the variance of mean decision purity are increasing over time, which is contradictory to out predictions.

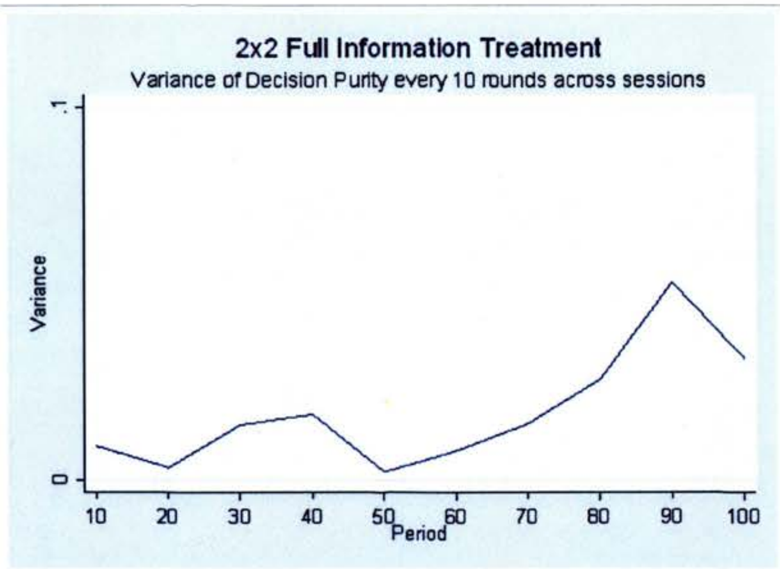

Figure 3.10: Variance of Average Putity in 2 × 2 Full Information Treatment

3.4.3 3×3 Aggregate Information Treatment

In Session 2, Group 1, we do not see any players converge to pure strategies. In Session 2, Group 2, we see only one player converge to a pure strategy, **B**. In Session 2, Group 3, we do not see any players converge to pure strategies. In Session 6, Group 1, we see only one player converge to a pure strategy, **C**. In Session 6, Group 2, we

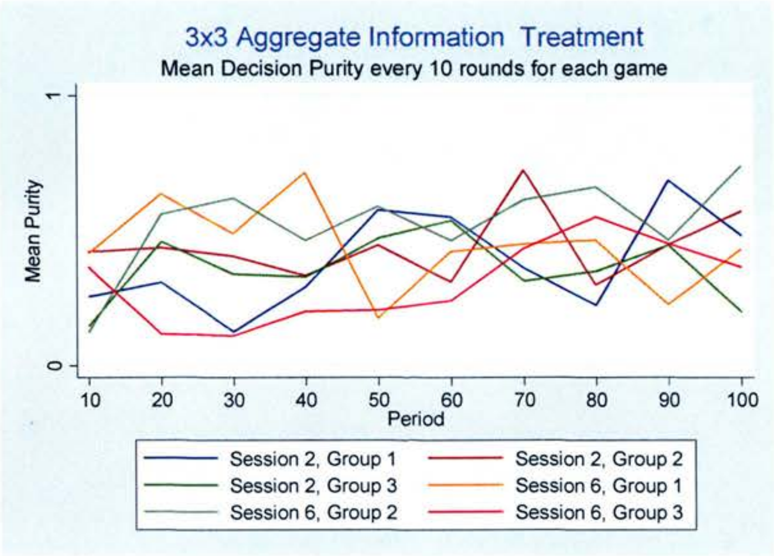

Figure 3.11: 3×3 Aggregate Information Treatment: Mean Decision Purity every 10 rounds for each game.

see two players converge to pure strategies, both to **B.** In Session 6, Group 3, we see only one player converge to a pure strategy, **B.**

In the six 3×3 finite population games with aggregate information, we see at least one player in half the games converges to pure strategy **B**.

In order to further establish evidence of convergence to pure strategies we now study the average purity across the treatment. In Figure 3.12, each data point represents that mean decision purity every 10 period across the whole treatment.

In the 3×3 aggregate information treatment sessions the mean decision purity increases over time with a coefficient of the linear regression model that is significant at

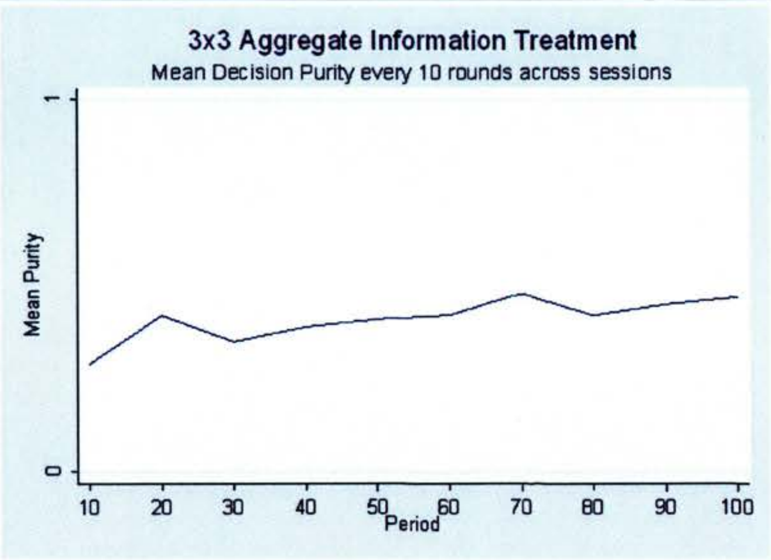

Figure 3.12: Average Putity in 3×3 Aggregate Information Treatment

the 1% level of significance. There is strong evidence to reject the null hypothesis that the coefficient is less than or equal to zero ($Ho : Coef <= 0 : p-value = 0.00283345$).

We also investigate the variance of mean decision purity every 10 rounds in each game (see Figure 3.13). However, there is little evidence to reject the null hypothesis that the variance of mean decision purity is increasing over time ($Ho : Coef >= 0 : p - value = 0.29923019$). Together, these two pieces of evidence provide encouraging support for our predictions that individual play should converge to pure strategies.

3.4.4 3×3 Full Information Treatment

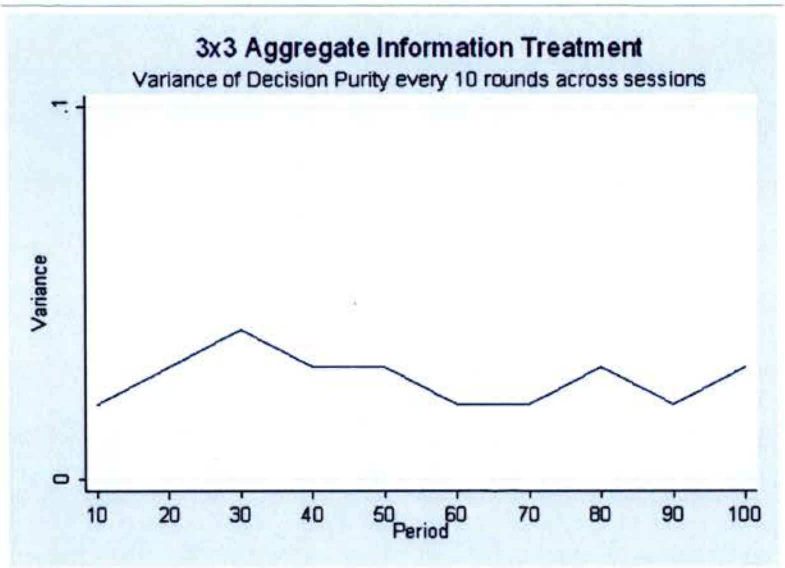

Figure 3.13: Variance of Average Putity in 3×3 Aggregate Information Treatment

In Session 4, Group 1, we do not see any players converge to pure strategies. In Session 4, Group 2, we only one player converge to a pure strategy, **B.** In Session 5, Group 1, we see two players converge to pure strategies, one to **A** and one **B**. In Session 5, Group 2, we see only one player converge to a pure strategy, **B**. In Session 5, Group 3, we see only one player converge to a pure strategy, **C.**

In the five 3×3 finite population games with full information, we see at least one player in over half the games converges to pure strategy **B**.

In order to further establish evidence of convergence to pure strategies we now study the average purity across the treatment. In Figure 3.15, each data point represents that mean decision purity every 10 period across the whole treatment.

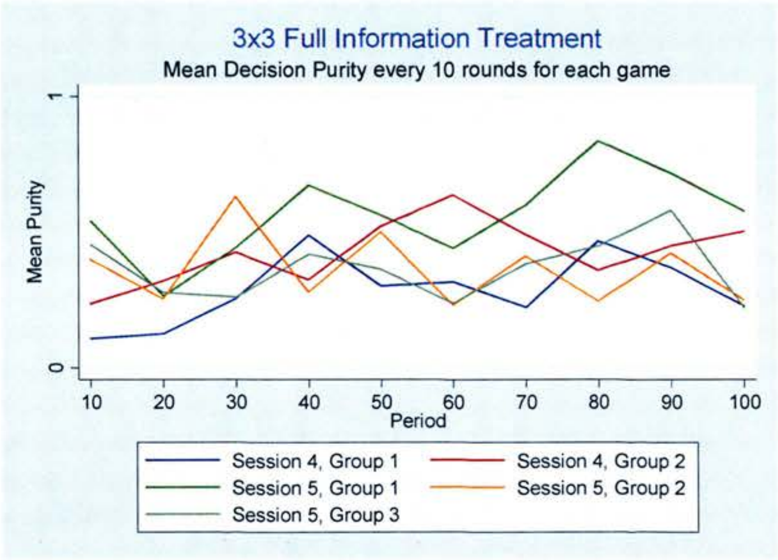

Figure 3.14: 3×3 Full Information Treatment: Mean Decision Purity every 10 rounds for each game.

Data from the 3×3 full information treatment sessions suggest similar behaviour as in the 3×3 aggregate information treatment sessions, both, which are not consistent with our predictions. In the 3×3 full information treatment sessions the mean decision purity increase over time with a coefficient of the linear regressing model that is significant at the 1% level of significance. In this treatment, there is strong evidence reject the null hypothesis that the coefficient is less than or equal to zero ($H_o : coef <= 0 : p - value = 0.06100645$).

We also investigate the variance of mean decision purity every 10 rounds in each game (see Figure 3.16). Again, there is no or little evidence against the null hypothesis that the variance of mean decision purity is increasing over time ($H_o : coef >= 0 :$

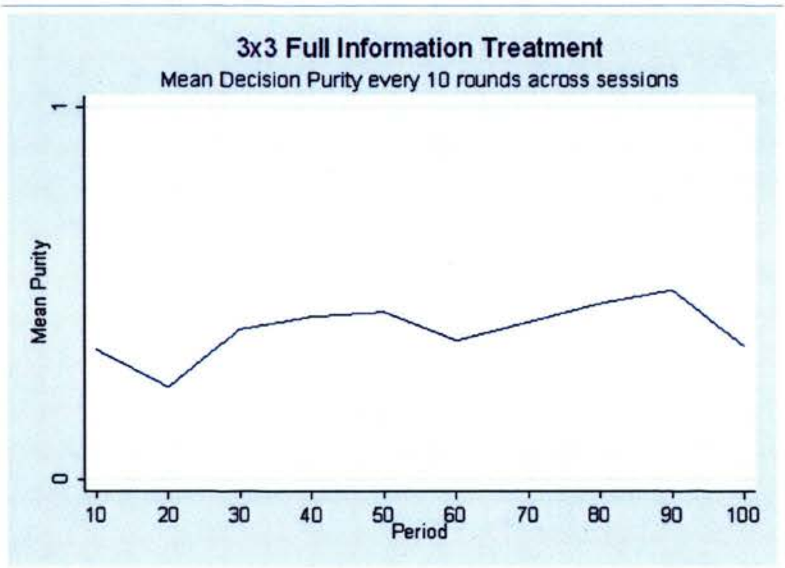

Figure 3.15: Average Putity in 3×3 Full Information Treatment

$p - value = 0.83160843$). Therefore, both the mean decision purity across treatment sessions and the variance of mean decision purity are increasing over time.

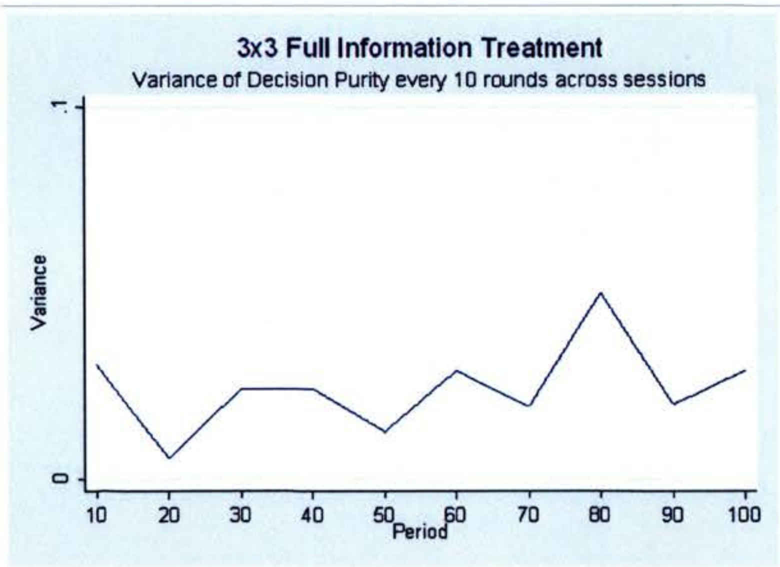

Figure 3.16: Variance of Average Putity in 3 × 3 Full Information Treatment

3.5 Discussion & Conclusions

Let us now consider our experimental results in light of our hypotheses set out in Section 3.2. First, we postulated that in all treatments long-run play should converge to the set of strict pure strategy equilibria. This hypothesis is derived from our results, which indicate that the long-run behaviour of boundedly rational agents in anti-coordination finite population games with random matching converges to a sorting outcome.

We believe that it is fair to say that the evidence from our experiments is less than conclusive on this issue. Although we did witness, in all anti-coordination finite population games run in the laboratory, some subjects who appeared to have settled on pure strategies, it is difficult to make any conclusive statements confirming that behaviour in the population converged to the set of strict pure strategy equilibria of the anti-coordination finite population game.

However, given the quality of the data, we feel there are opportunities for further analysis. There is ample evidence in the literature that, for games with multiple Nash equilibria, individual behaviour observed in the laboratory is often not consistent with any of them, while at the same time, on aggregate, behaviour is close to the symmetric mixed strategy Nash equilibrium. Our experimental results confirm this observation.

With this in mind, it worth noting that it would be possible to study the non-equilibrium dynamics of play in our anti-coordination finite population games as they also admit multiple equilibria. Convergence may have been slowed by the complexity of the strategic environment or noise in the population.[5] Genin and Katok (2006) suggest that if non-equilibrium dynamics are considered, it can be shown that aggregate behaviour is consistent with behaviour observed in the laboratory. In this case it would be useful to measure the distance the actual play is from other Nash equilibria, specifically the asymmetric mixed strategy Nash equilibria.

Our second hypothesis postulated that convergence to the predicted outcomes would happen faster in the full information treatments than in the aggregate

[5]In our design we did try to counteract this issue by choosing payoff matrices that admitted strict pure strategy Nash equilibria for a minimal population size.

information treatments. Our experimental results do not provide any positive evidence for this prediction. It should be noted that this has been the experience in experiments of other games and, since there are no explicit models which provide clear intuitions, it is difficult to make any progress in this direction.

Regardless, we believe that it would be instructive to further study the response dynamics to the information flow. In our treatments we postulated that subjects responded to payoffs; this is most likely the case for the aggregate information treatments. However, the extra feedback information could possibly be effecting the strength of the payoff reinforcement mechanism as participants respond to other players' behaviour instead.

An alternative hypothesis incorporating a reinforcement feedback mechanism might provide some intuition to our experimental data. Recall that the theoretical predictions obtained from studying the long-run behaviour of boundedly rational agents, learning in accordance with a reinforcement learning model, in anti-coordination finite population games, implies that people tend to minimise bad experiences and maximise good ones. In fact this is exactly what is assumed by the Erev and Roth (1998) reinforcement learning model.

The idea is that participants behave in accordance with the tenants of a reinforcement mechanism, although in a more heuristic manner. Note that in both game treatments, and assuming the opponent is randomising their choice over all the alternatives, the option with the highest expected payoff is **B**. It is a striking observation that where there were participants appearing to converge to a pure strategy, it was, more often than not, to option **B**. It is quite possible that this behaviour constitutes part of the story behind the non-equilibrium dynamics in anti-coordination finite population games with random matching.

The purpose of this study was to employ the theory of learning in games to further understand individual decision making in population games. We find, as previous experimental studies have remarked, that, on aggregate, behaviour is close to the symmetric mixed strategy Nash equilibrium predicted by evolutionary game theory. Furthermore, we have produced experimental data that reasserts Friedman's (1996) observation which originally motivated our study:

"...it appears that some players play 'Hawk', some play 'Dove' and others switch back and forth."

Appendix 3.A Experimental Data

3.A.1 2×2 Aggregate Information Treatment Data

Session 1, Group 1

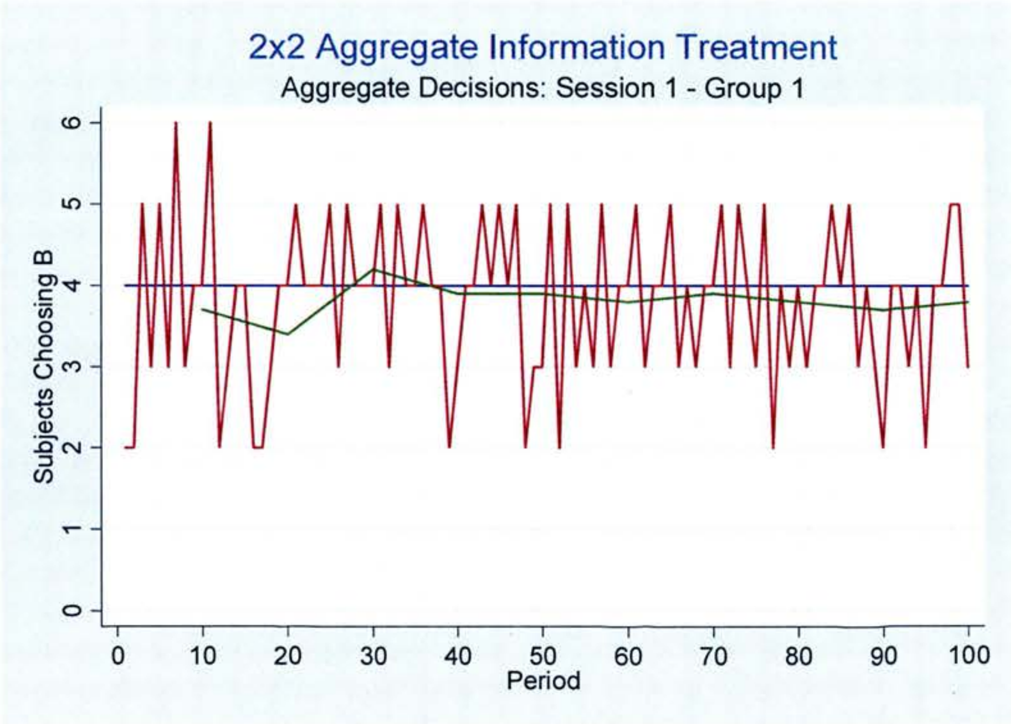

Figure 3.17: 2 × 2 Aggregate Information Treatment (Session 1, Group 1): Aggregate Decisions

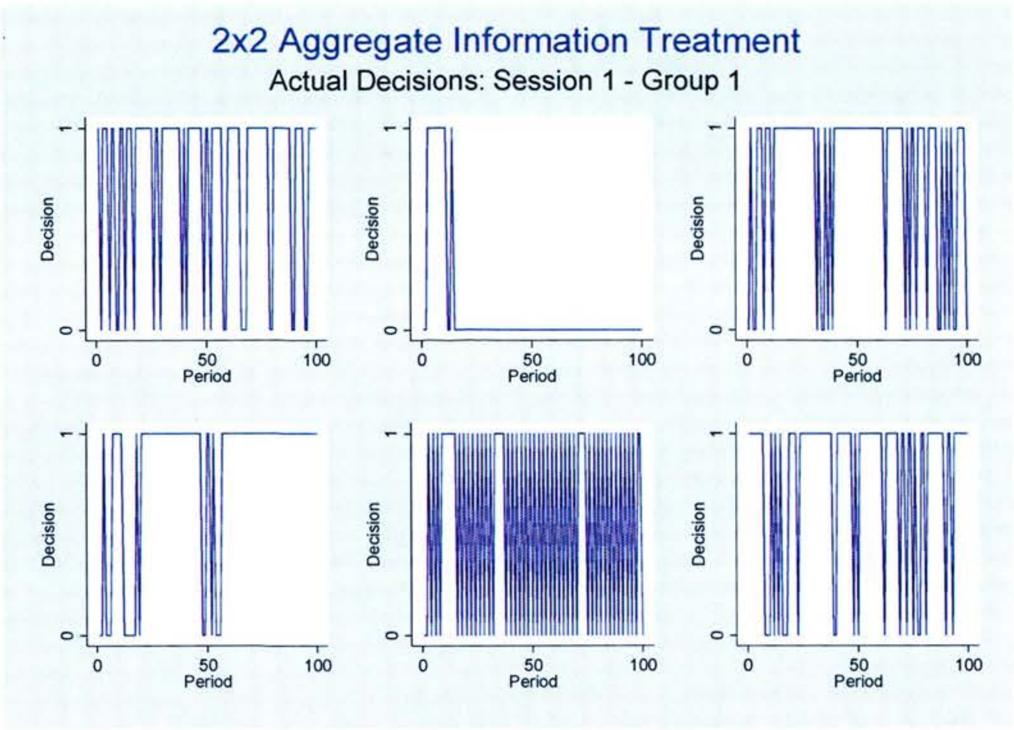

Figure 3.18: 2×2 Aggregate Information Treatment (Session 1, Group 1): Individual Decisions

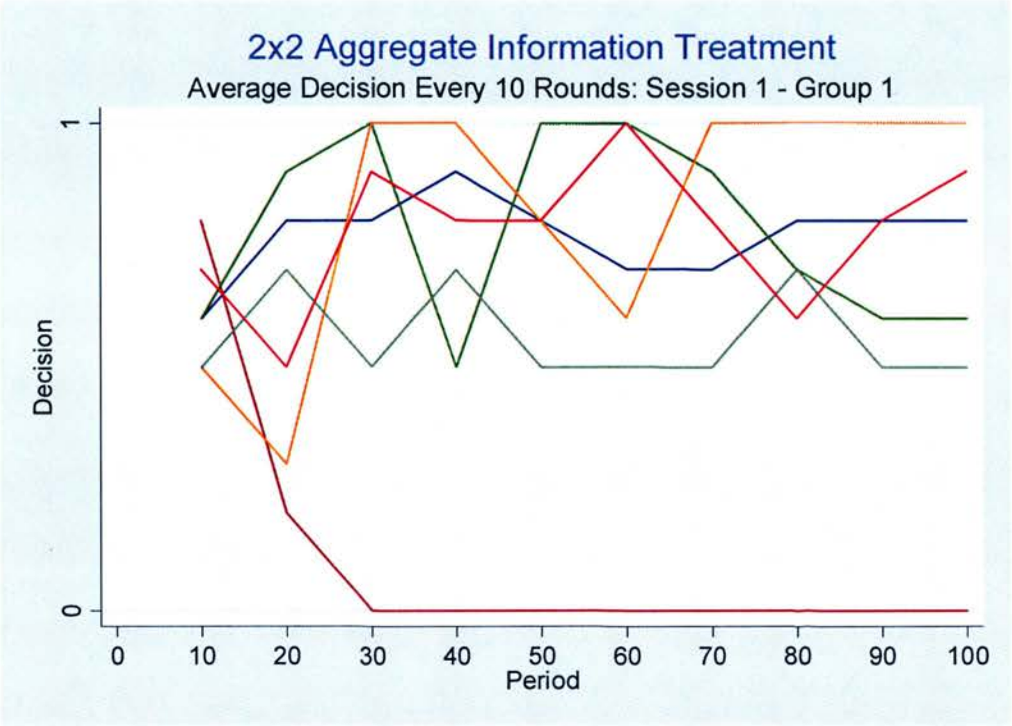

Figure 3.19: 2 × 2 Aggregate Information Treatment (Session 1, Group 1): Average Individual Decision Every 10 Rounds

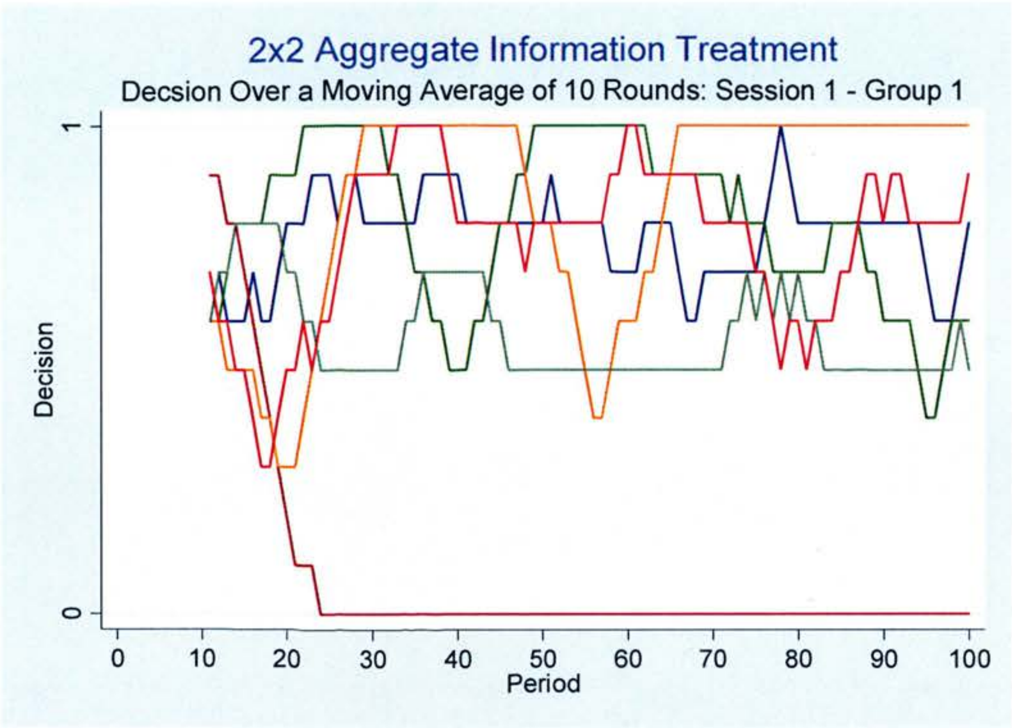

Figure 3.20: 2 × 2 Aggregate Information Treatment (Session 1, Group 1): Average Individual Decision Over Moving 10 Rounds

Session 1, Group 2

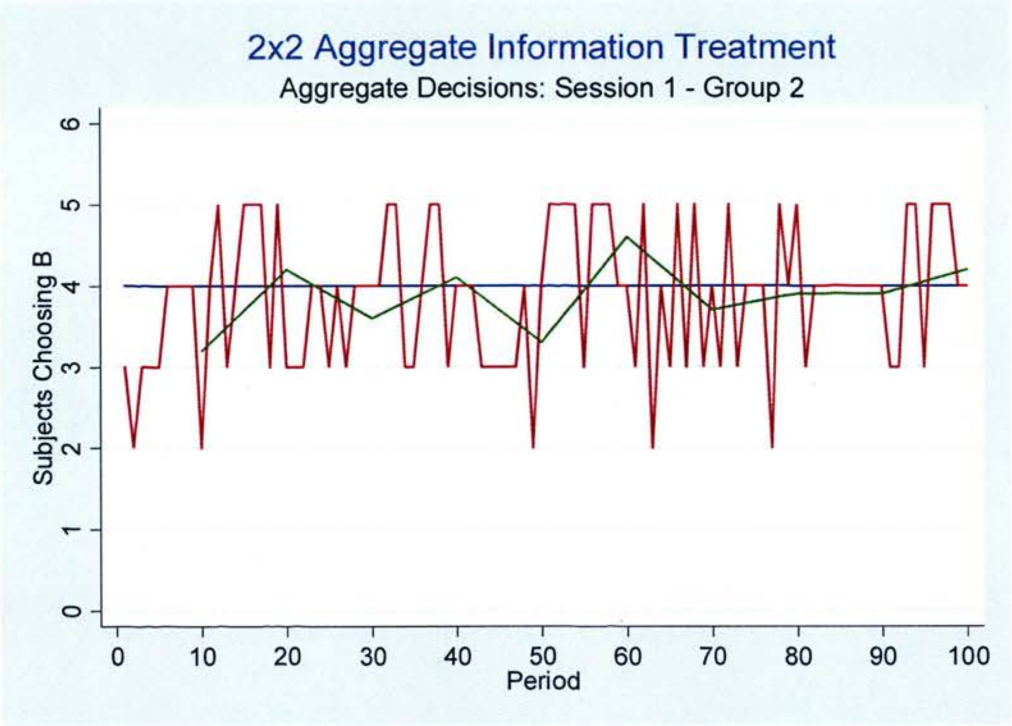

Figure 3.21: 2×2 Aggregate Information Treatment (Session 1, Group 2): Aggregate Decisions

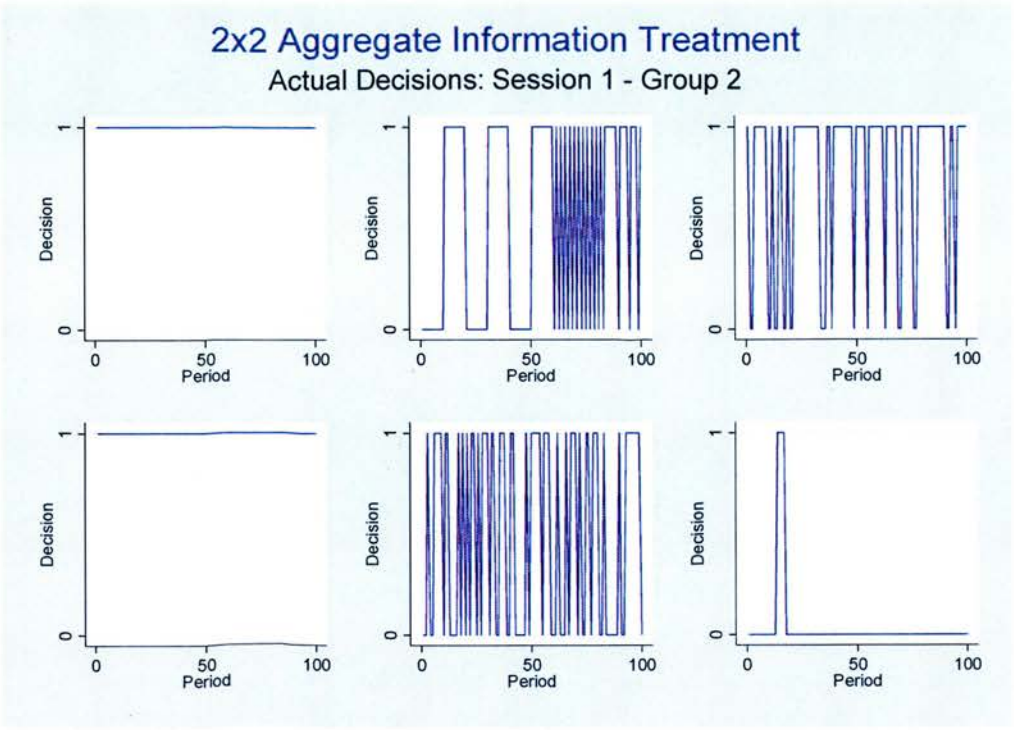

Figure 3.22: 2 × 2 Aggregate Information Treatment (Session 1, Group 2): Individual Decisions

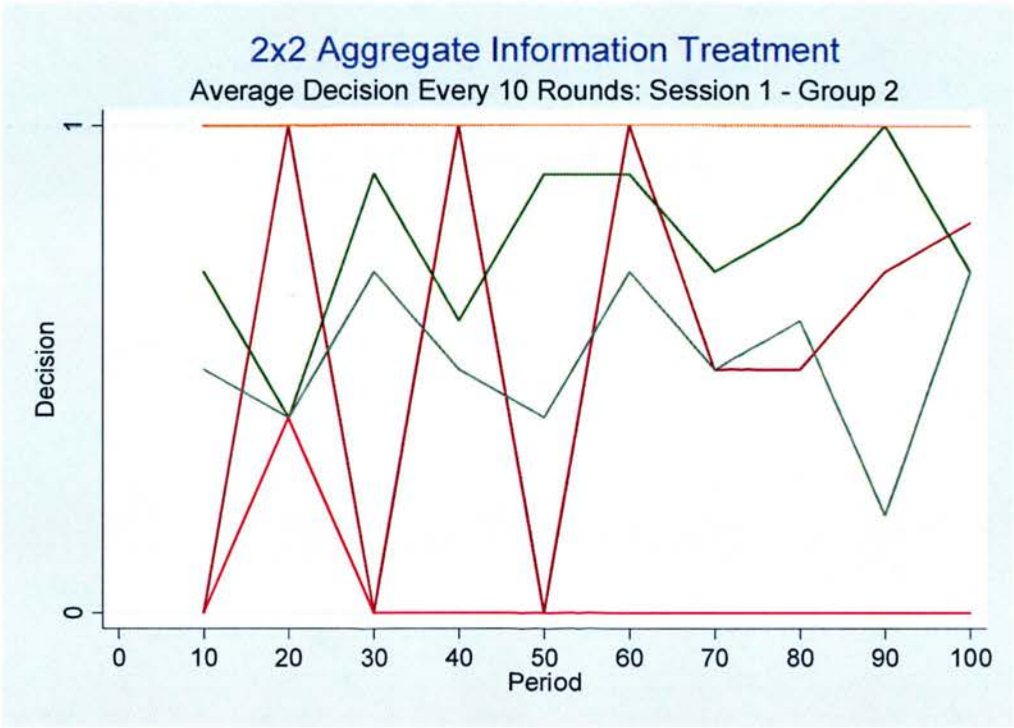

Figure 3.23: 2×2 Aggregate Information Treatment (Session 1, Group 2)

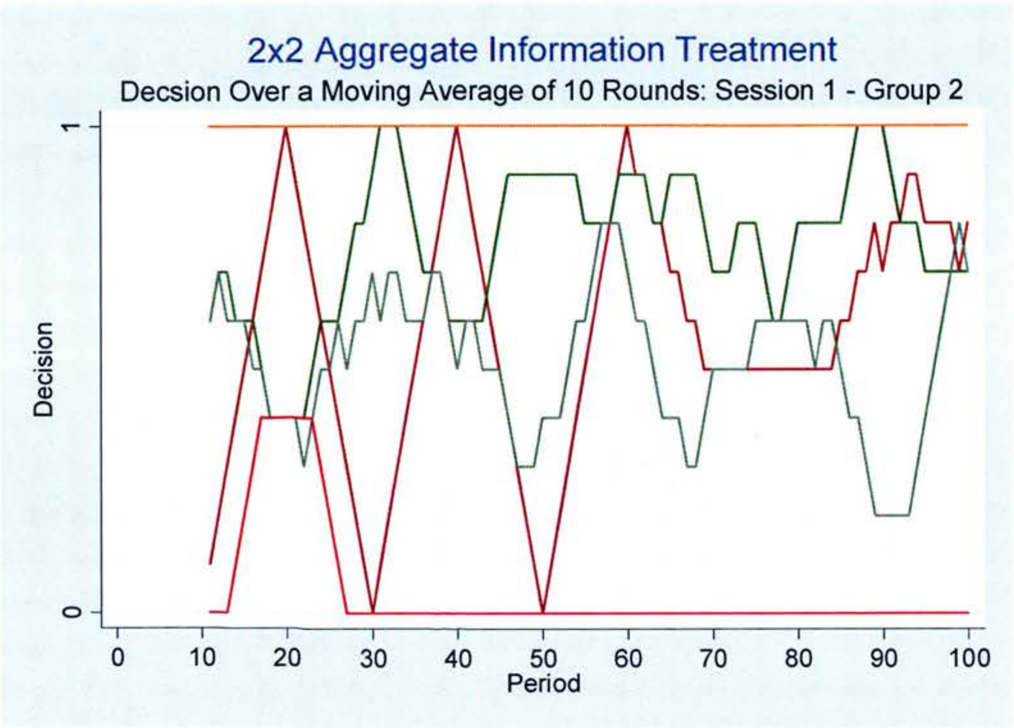

Figure 3.24: 2 × 2 Aggregate Information Treatment (Session 1, Group 2): Average Individual Decision Over Moving 10 Rounds

Session 1, Group 3

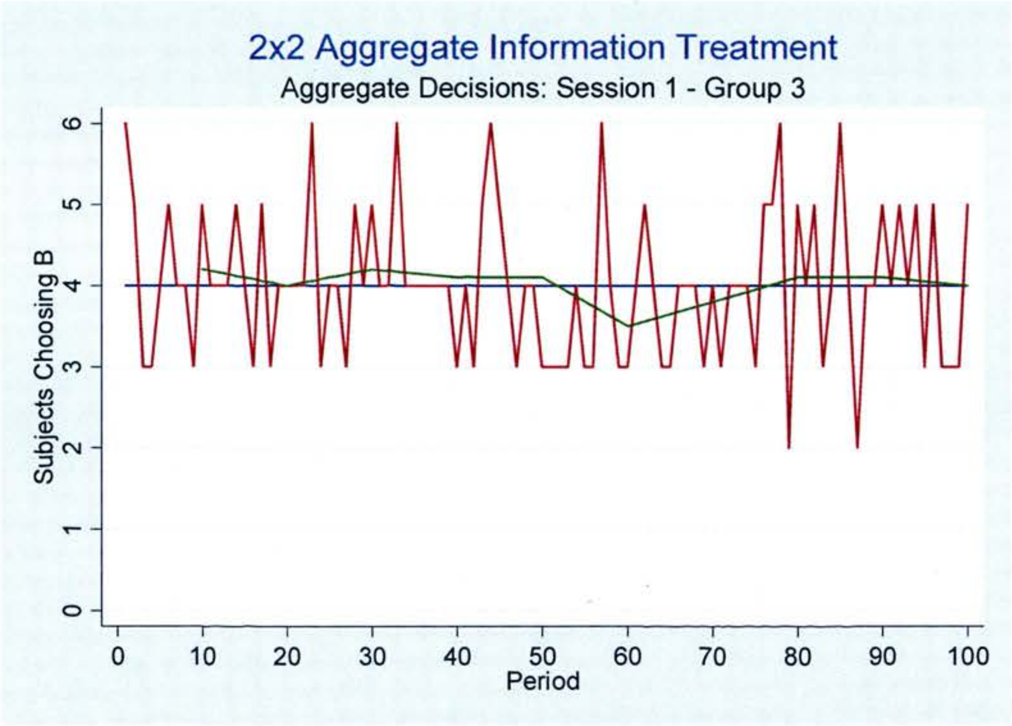

Figure 3.25: 2×2 Aggregate Information Treatment (Session 1, Group 3): Aggregate Decisions

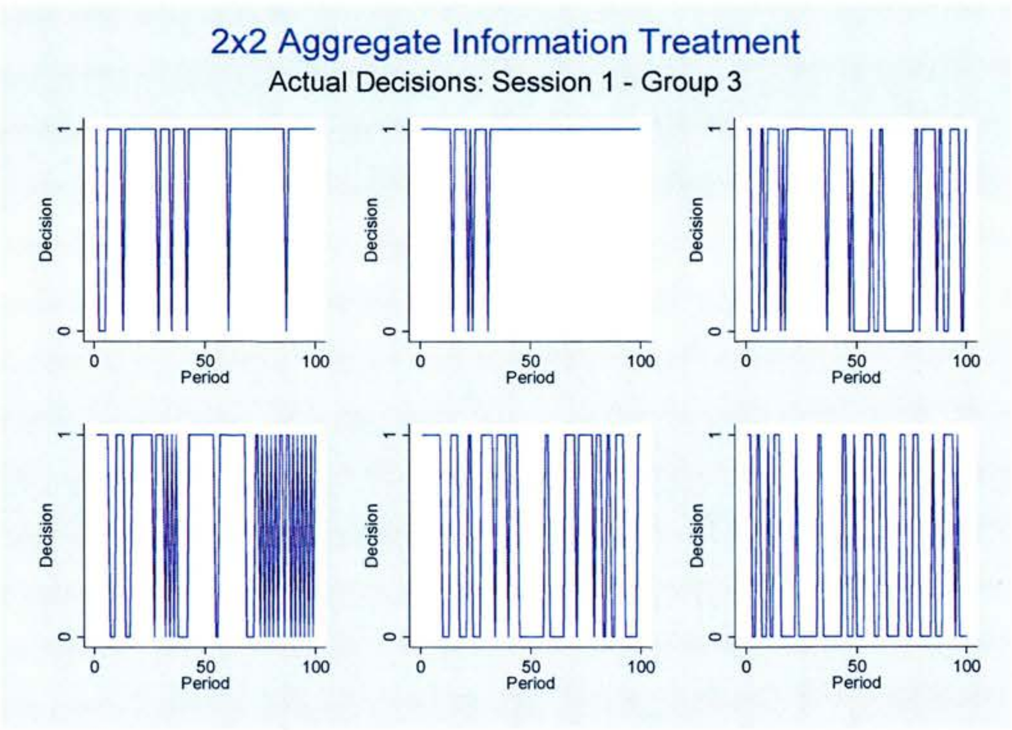

Figure 3.26: 2×2 Aggregate Information Treatment (Session 1, Group 3): Individual Decisions

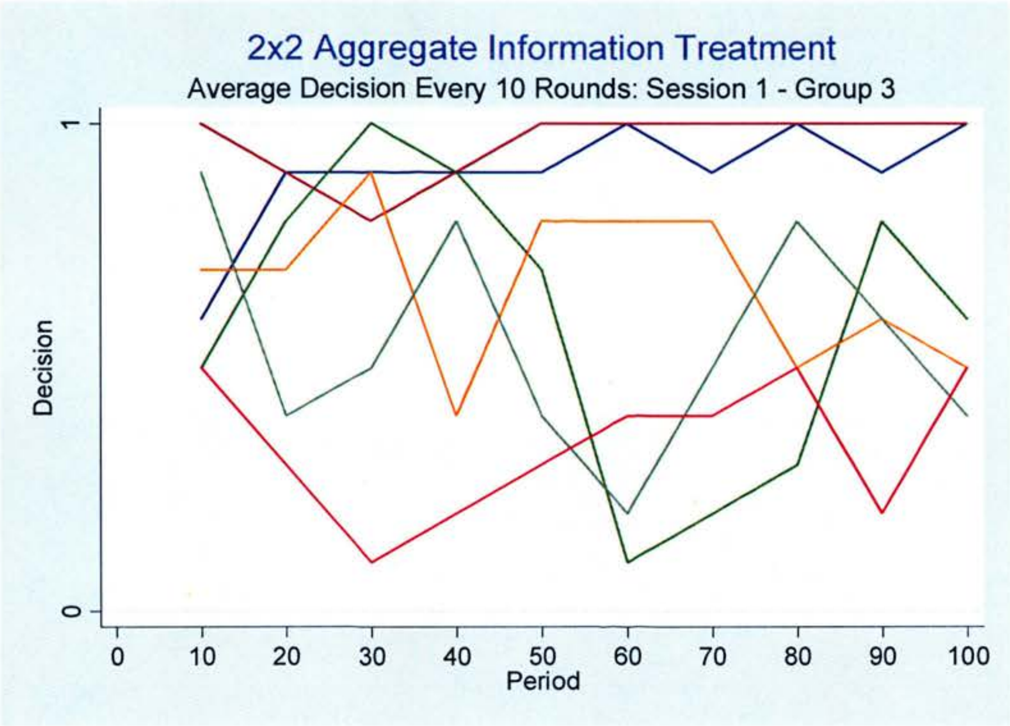

Figure 3.27: 2×2 Aggregate Information Treatment (Session 1, Group 3)

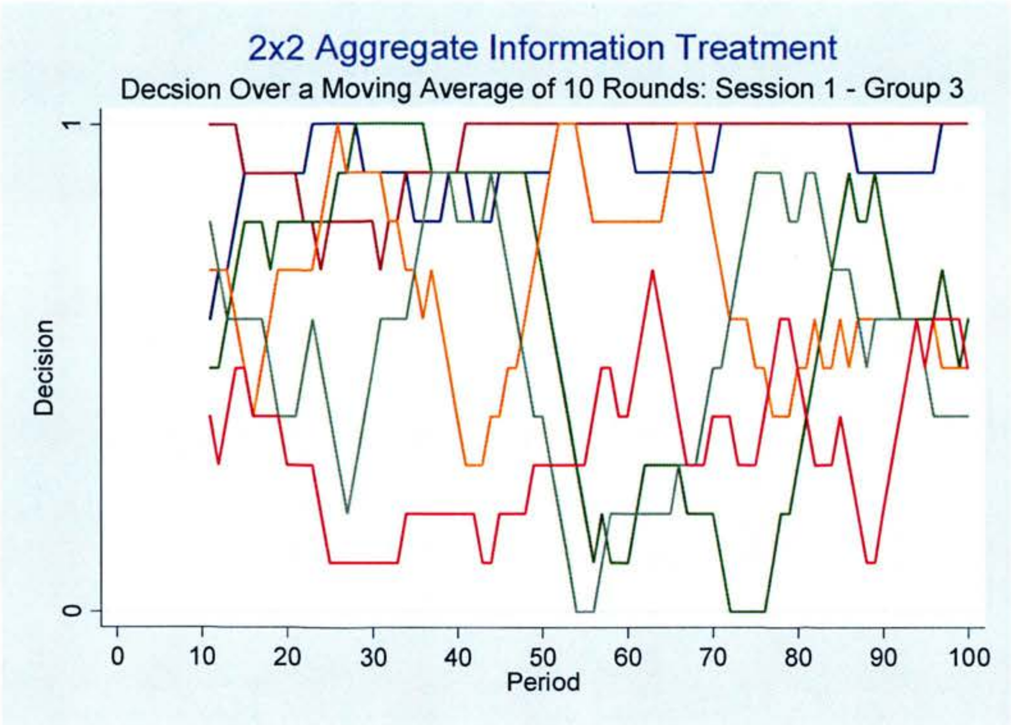

Figure 3.28: 2 × 2 Aggregate Information Treatment (Session 1, Group 3): Average Individual Decision Over Moving 10 Rounds

Session 7, Group 1

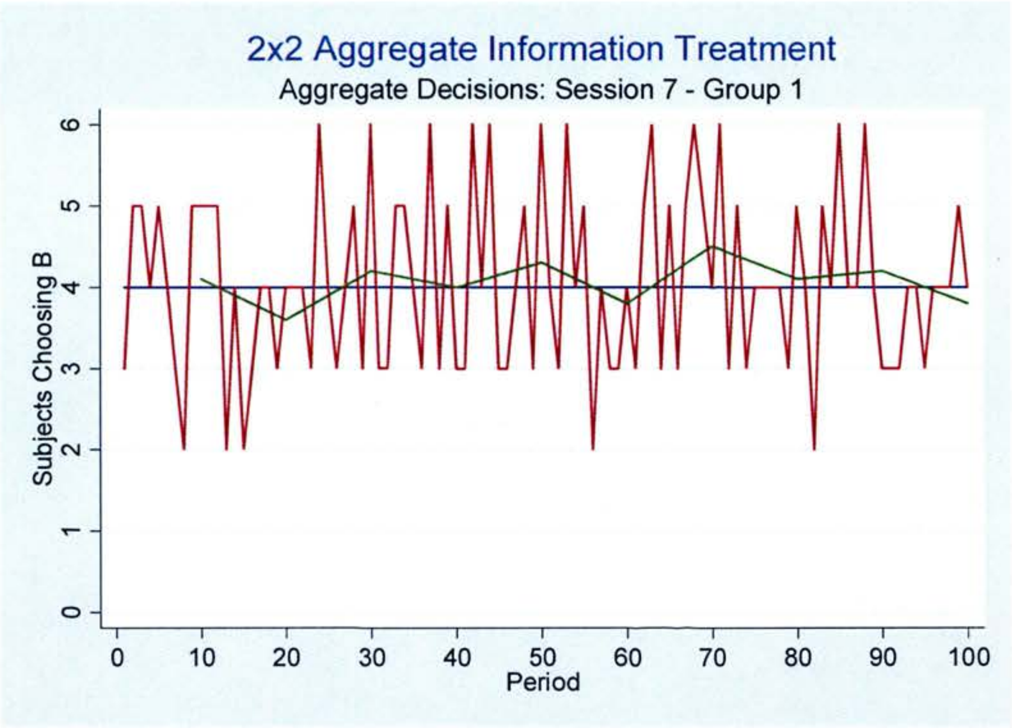

Figure 3.29: 2×2 Aggregate Information Treatment (Session 7, Group 1): Aggregate Decisions

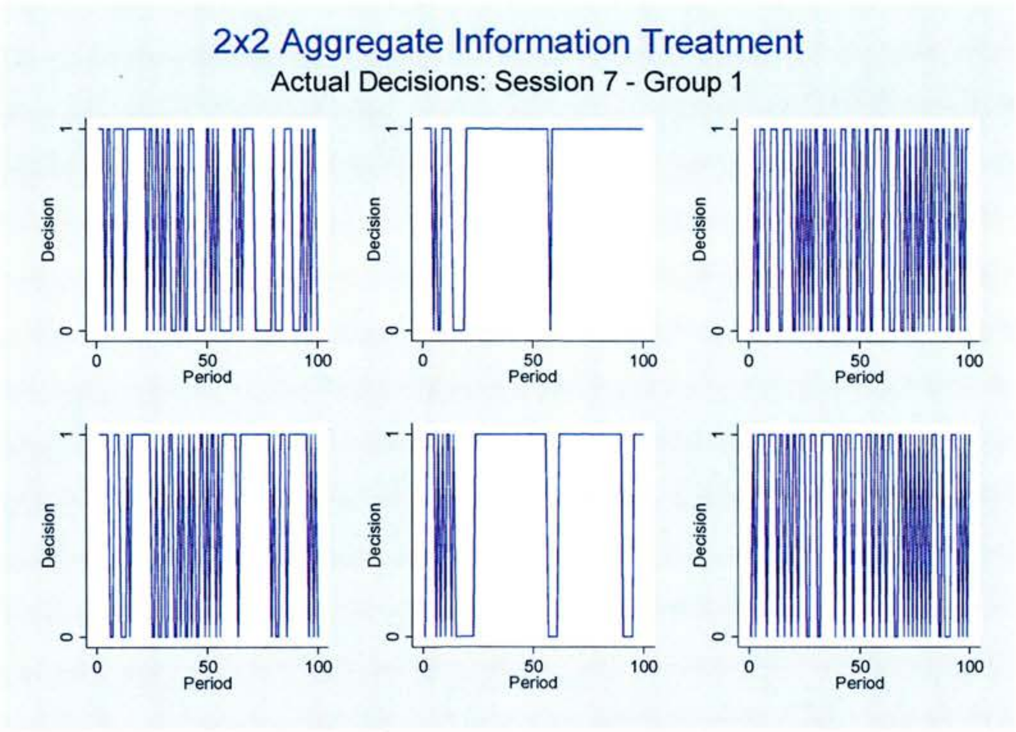

Figure 3.30: 2×2 Aggregate Information Treatment (Session 7, Group 1): Individual Decisions

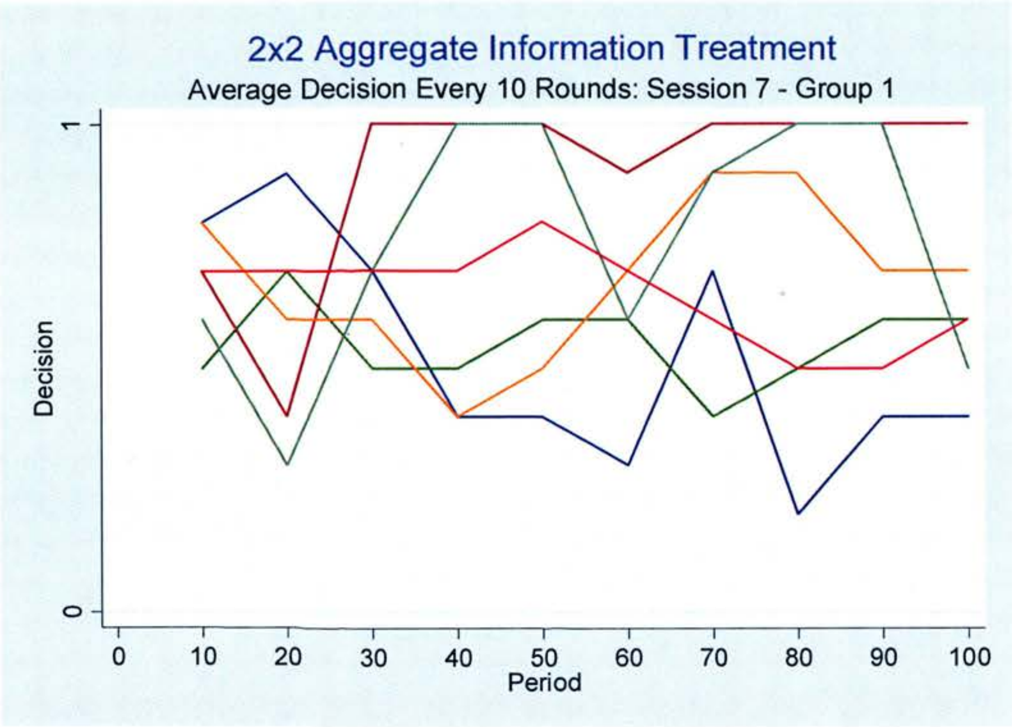

Figure 3.31: 2 × 2 Aggregate Information Treatment (Session 7, Group 1): Average Individual Decision Every 10 Rounds

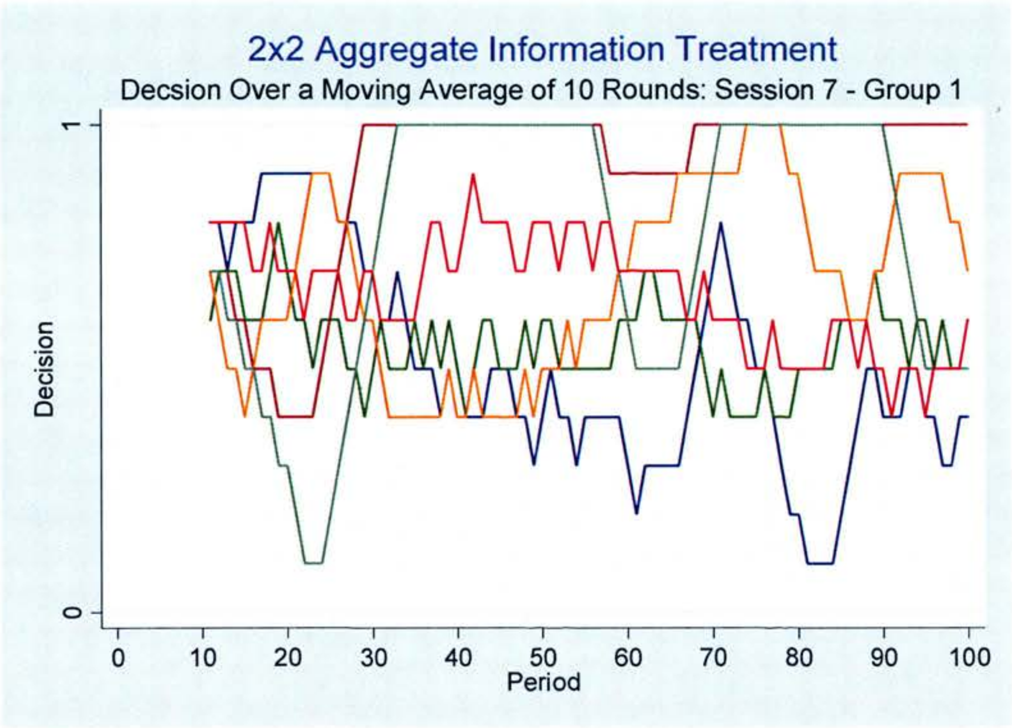

Figure 3.32: 2 × 2 Aggregate Information Treatment (Session 7, Group 1): Average Individual Decision Over Moving 10 Rounds

Session 7, Group 2

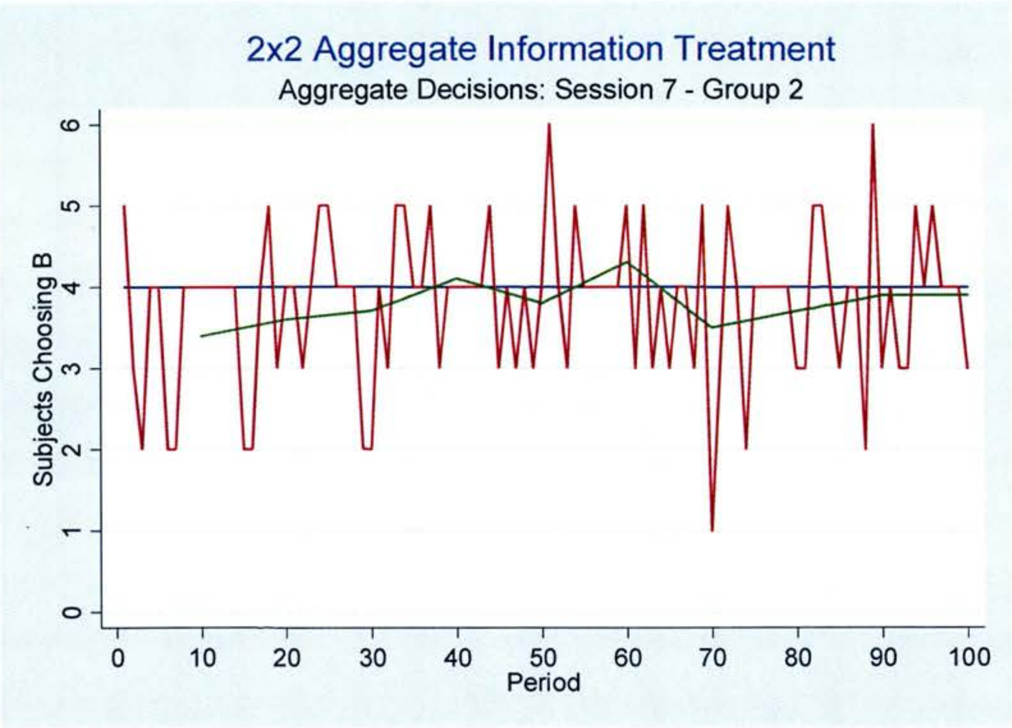

Figure 3.33: 2×2 Aggregate Information Treatment (Session 7, Group 2): Aggregate Decisions

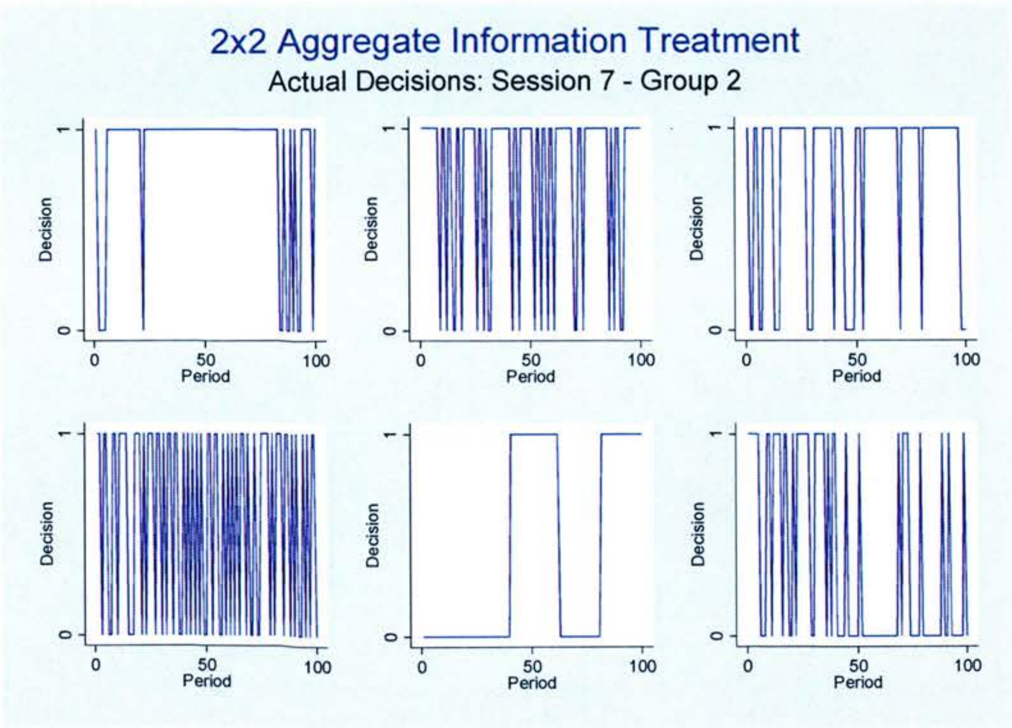

Figure 3.34: 2 × 2 Aggregate Information Treatment (Session 7, Group 2): Individual Decisions

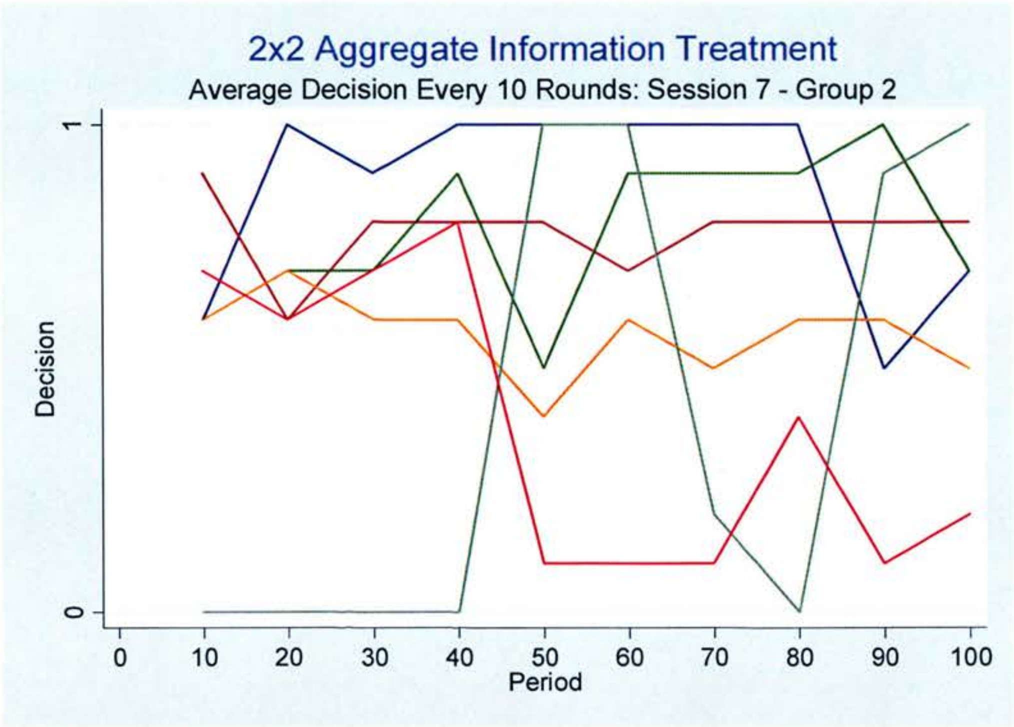

Figure 3.35: 2 × 2 Aggregate Information Treatment (Session 7, Group 2): Average Individual Decision Every 10 Rounds

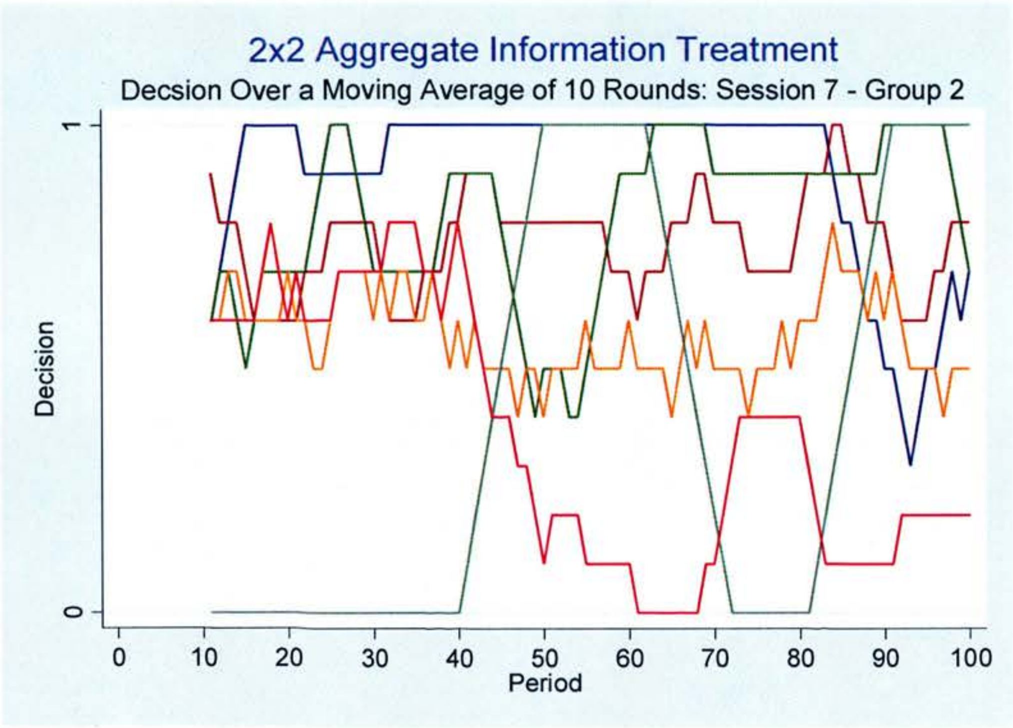

Figure 3.36: 2 × 2 Aggregate Information Treatment (Session 7, Group 2): Average Individual Decision Over Moving 10 Rounds

Session 7, Group 3

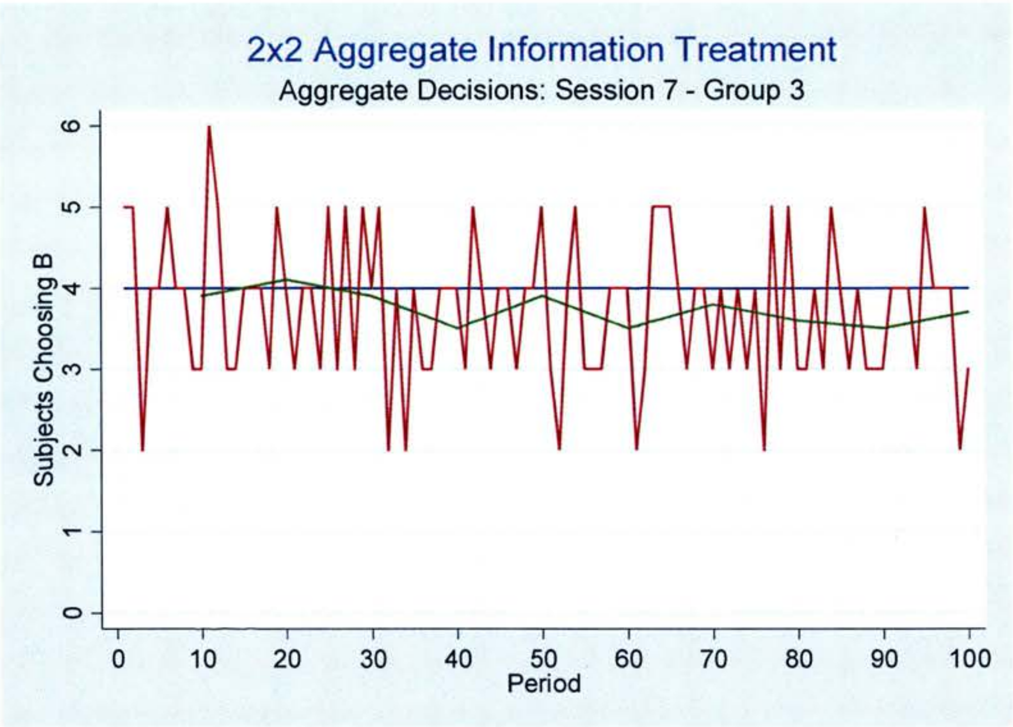

Figure 3.37: 2 × 2 Aggregate Information Treatment (Session 7, Group 3): Aggregate Decisions

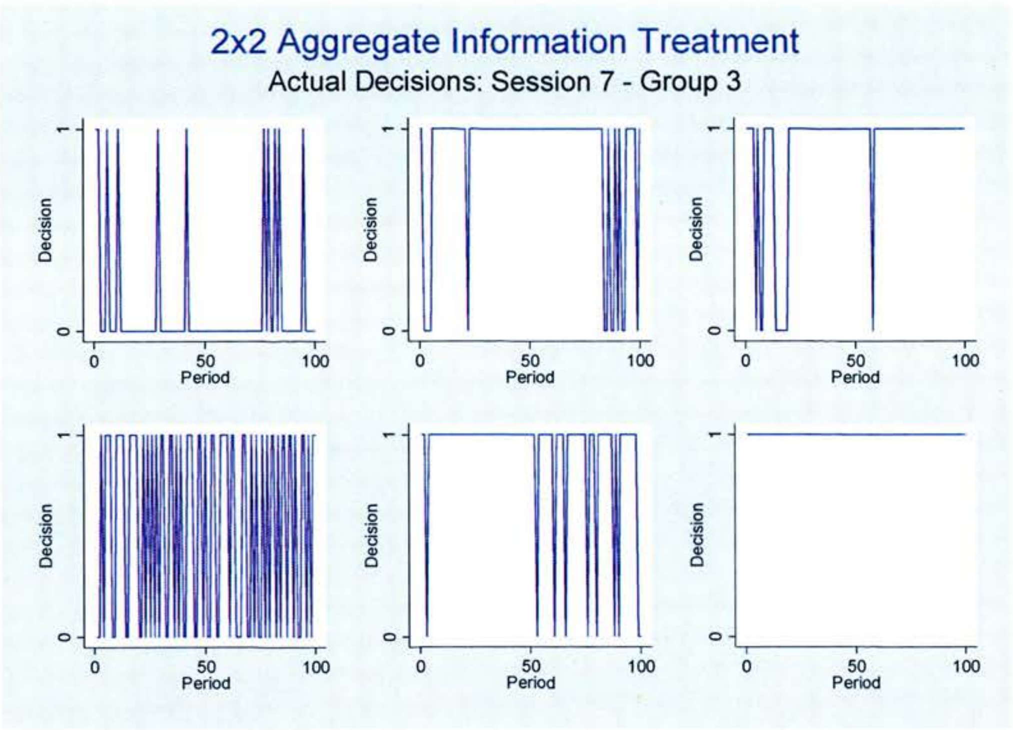

Figure 3.38: 2 × 2 Aggregate Information Treatment (Session 7, Group 3): Individual Decisions

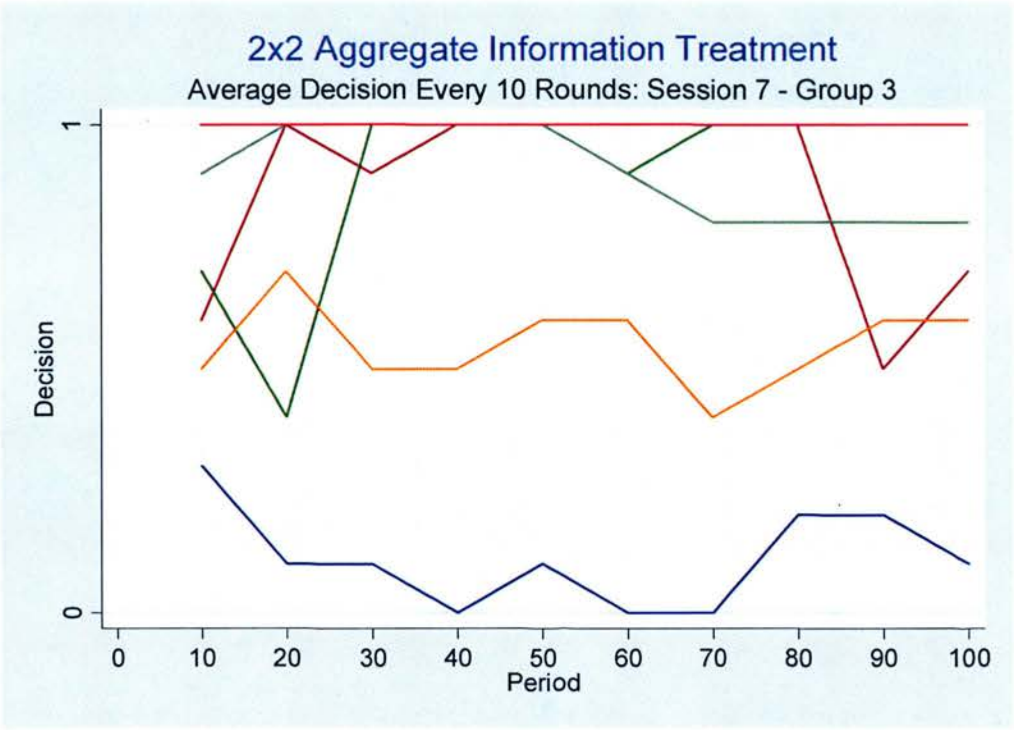

Figure 3.39: 2 × 2 Aggregate Information Treatment (Session 7, Group 3): Average Individual Decision Every 10 Rounds

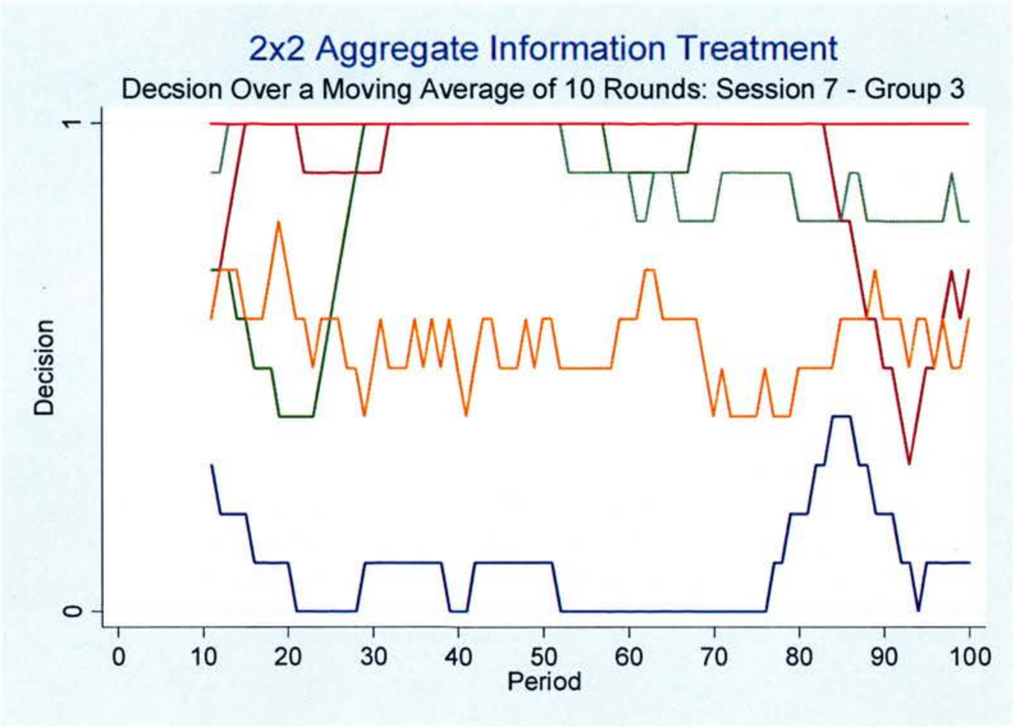

Figure 3.40: 2×2 Aggregate Information Treatment (Session 7, Group 3): Average Individual Decision Over Moving 10 Rounds

3.A.2 2×2 Full Information Treatment Data

Session 3, Group 1

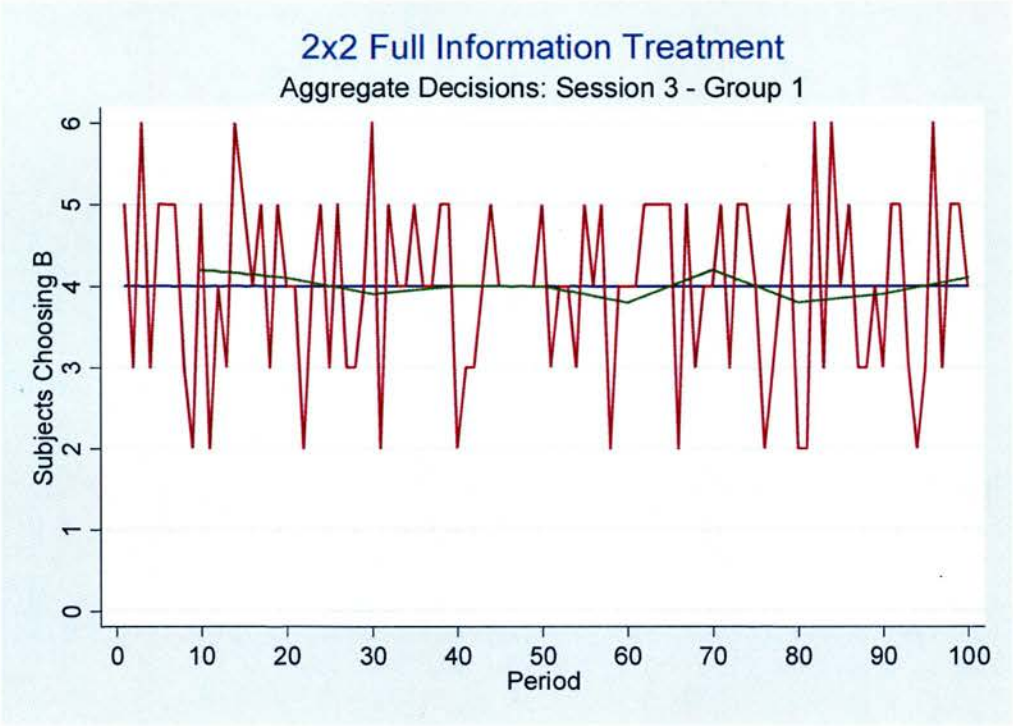

Figure 3.41: 2 × 2 Full Information Treatment (Session 3, Group 1): Aggregate Decisions

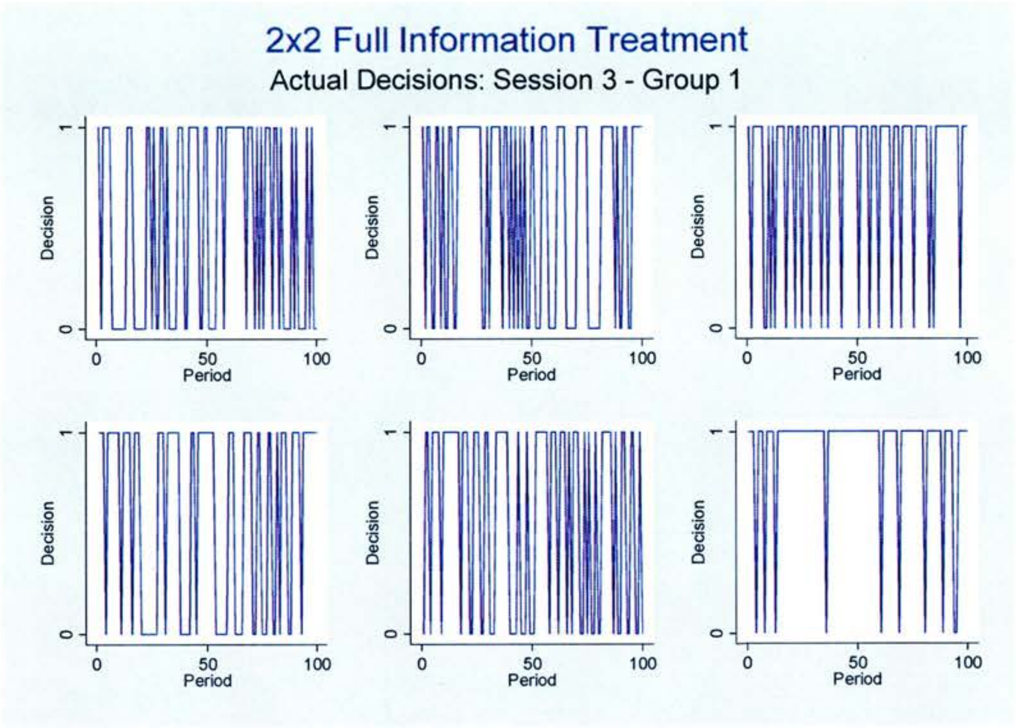

Figure 3.42: 2 × 2 Full Information Treatment (Session 3, Group 1): Individual Decisions

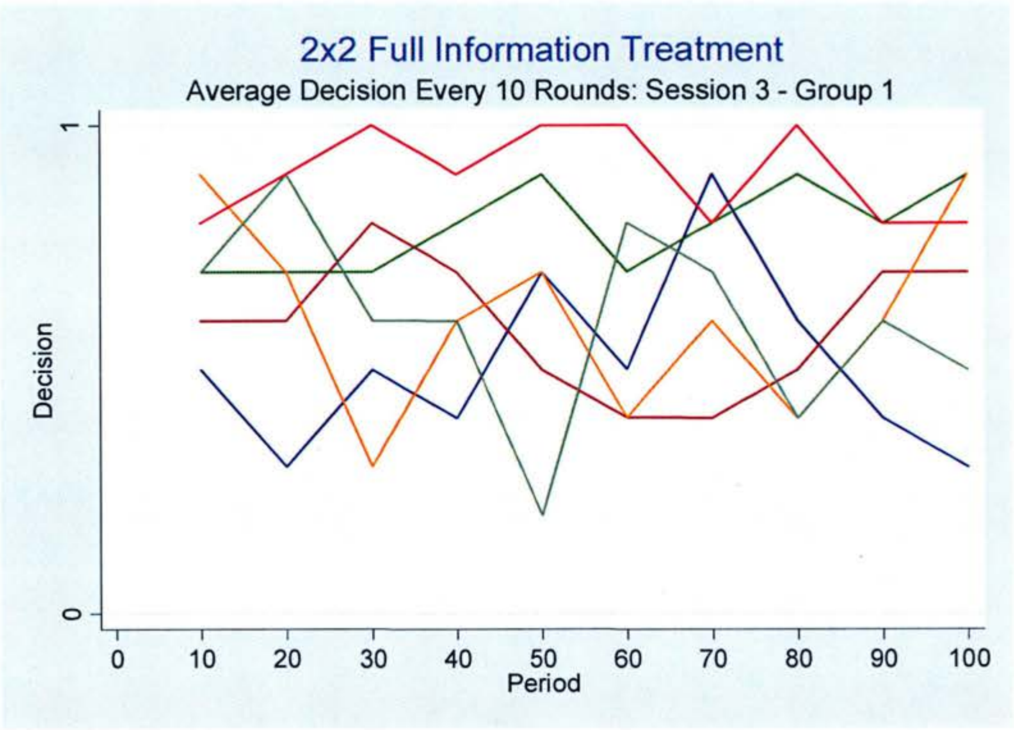

Figure 3.43: 2 × 2 Full Information Treatment (Session 3, Group 1): Average Individual Decision Every 10 Rounds

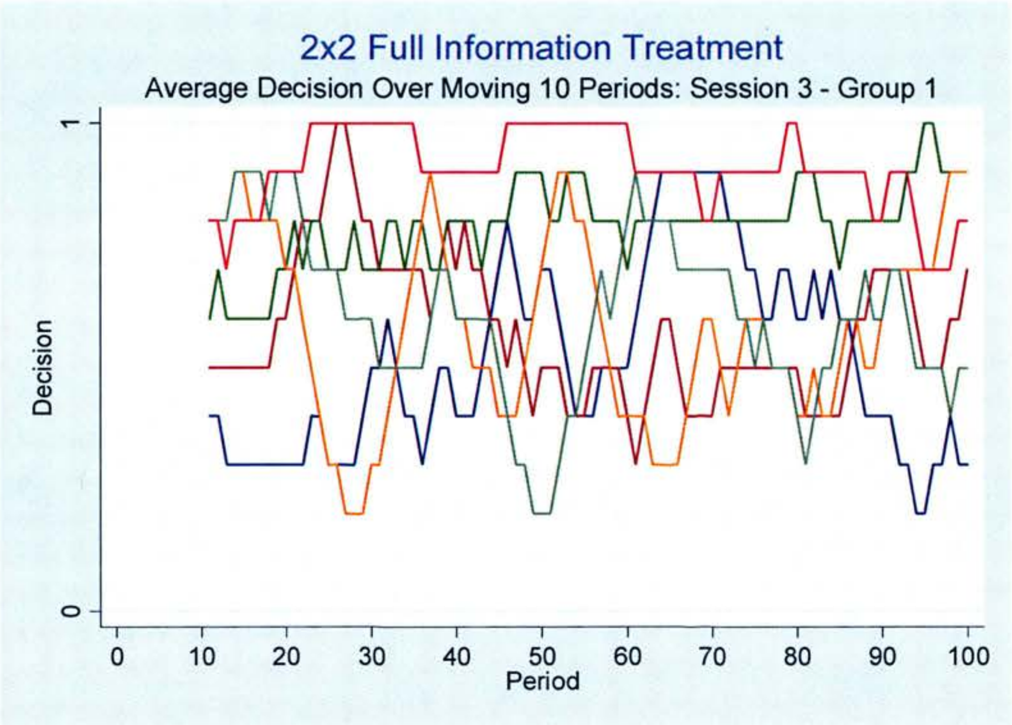

Figure 3.44: 2 × 2 Full Information Treatment (Session 3, Group 1): Average Individual Decision Over Moving 10 Rounds

Session 3, Group 2

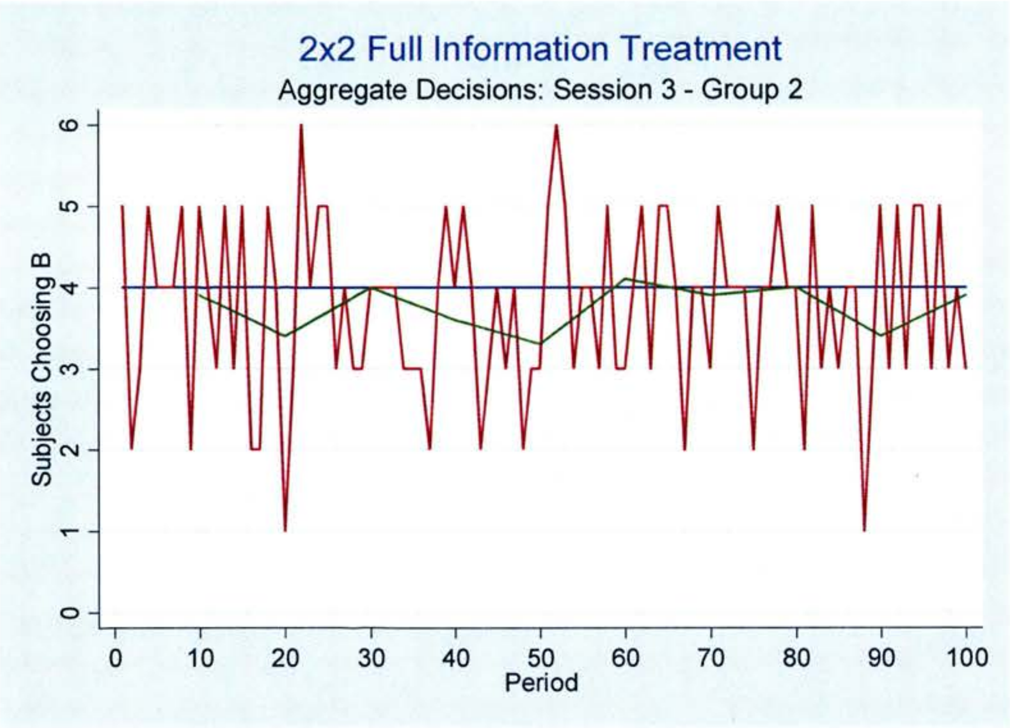

Figure 3.45: 2 × 2 Full Information Treatment (Session 3, Group 2): Aggregate Decisions

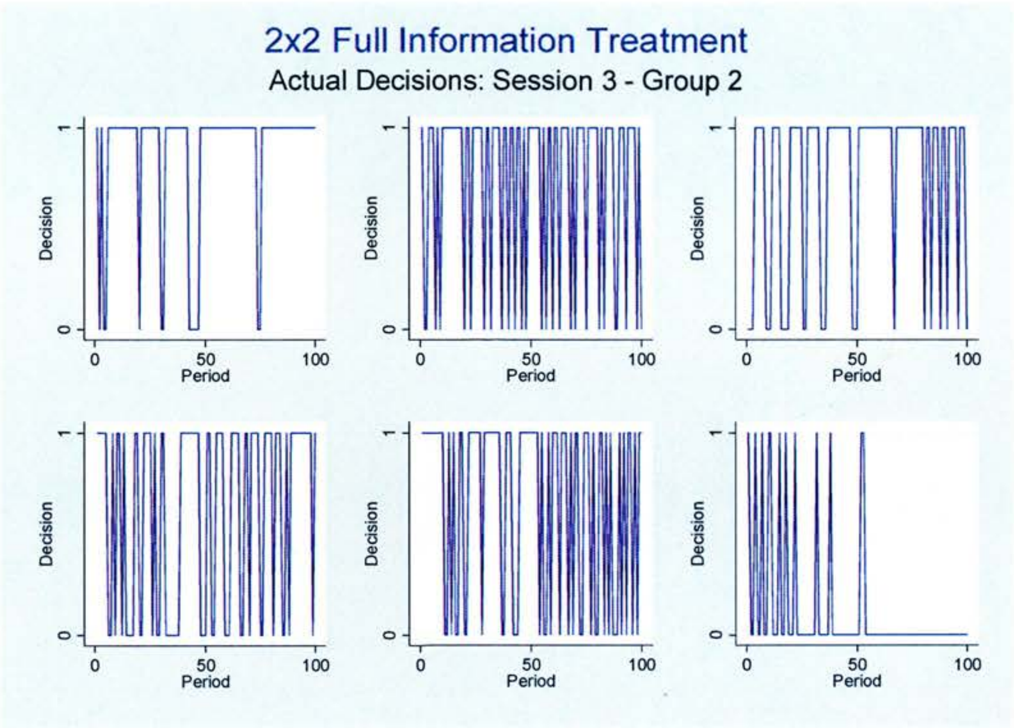

Figure 3.46: 2 × 2 Full Information Treatment (Session 3, Group 2): Individual Decisions

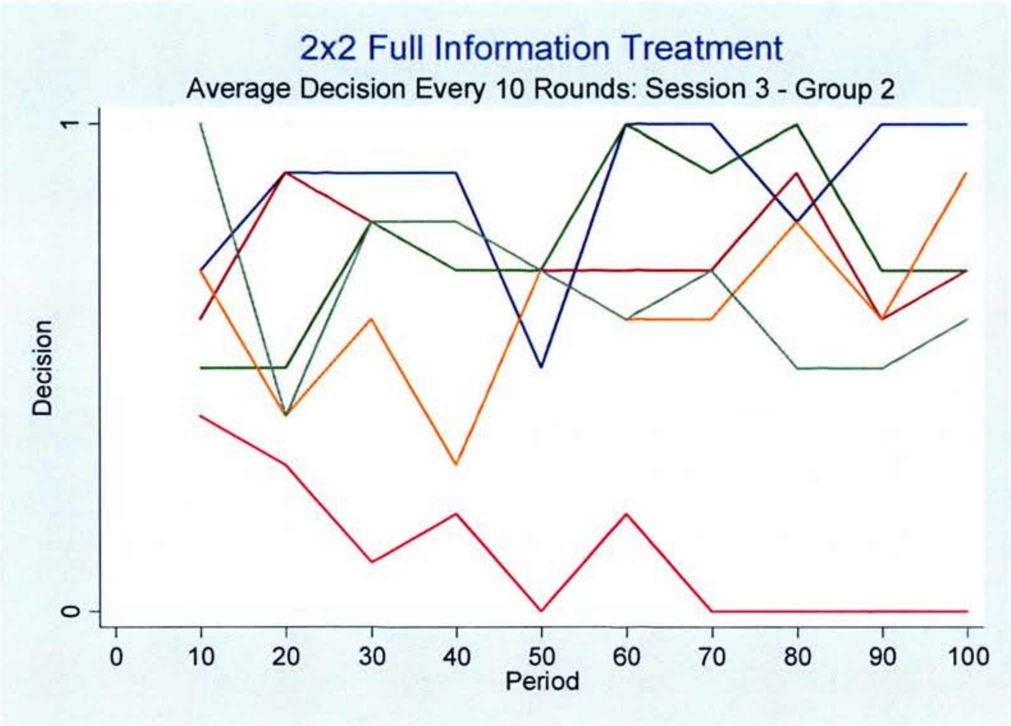

Figure 3.47: 2 × 2 Full Information Treatment (Session 3, Group 2): Average Individual Decision Every 10 Rounds

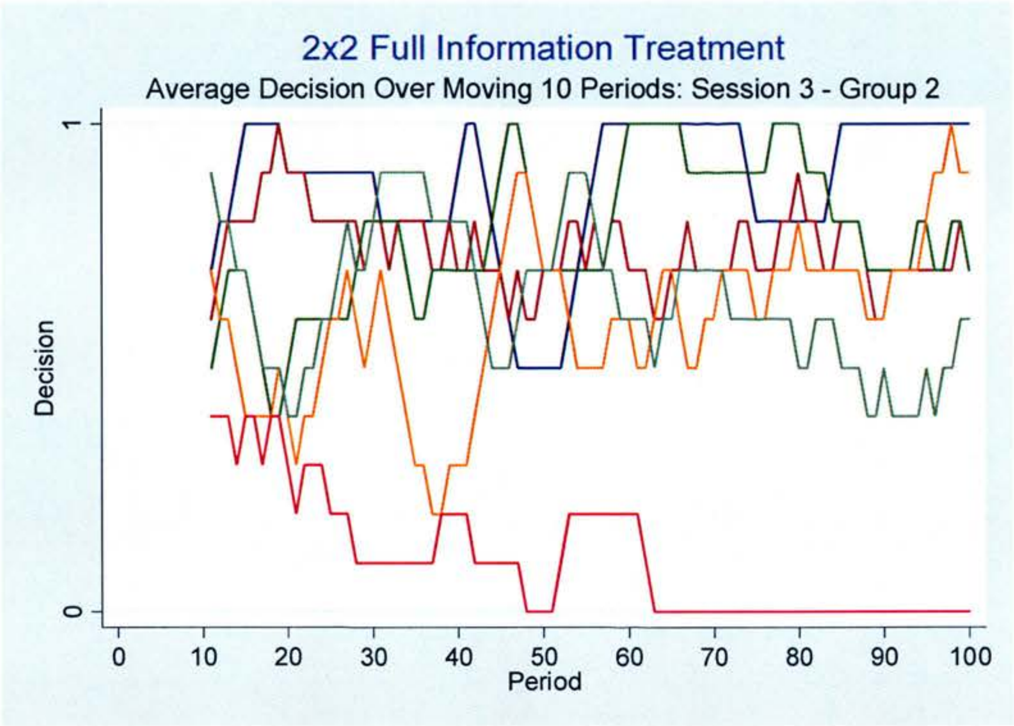

Figure 3.48: 2 × 2 Full Information Treatment (Session 3, Group 2): Average Individual Decision Over Moving 10 Rounds

Session 3, Group 3

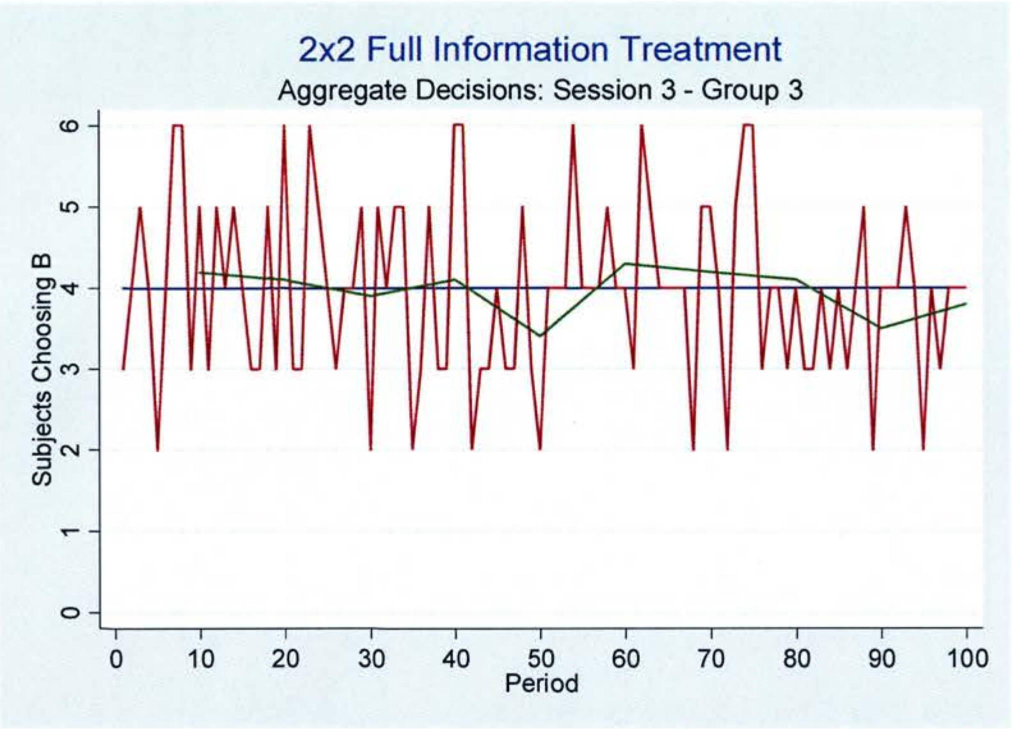

Figure 3.49: 2 × 2 Full Information Treatment (Session 3, Group 3): Aggregate Decisions

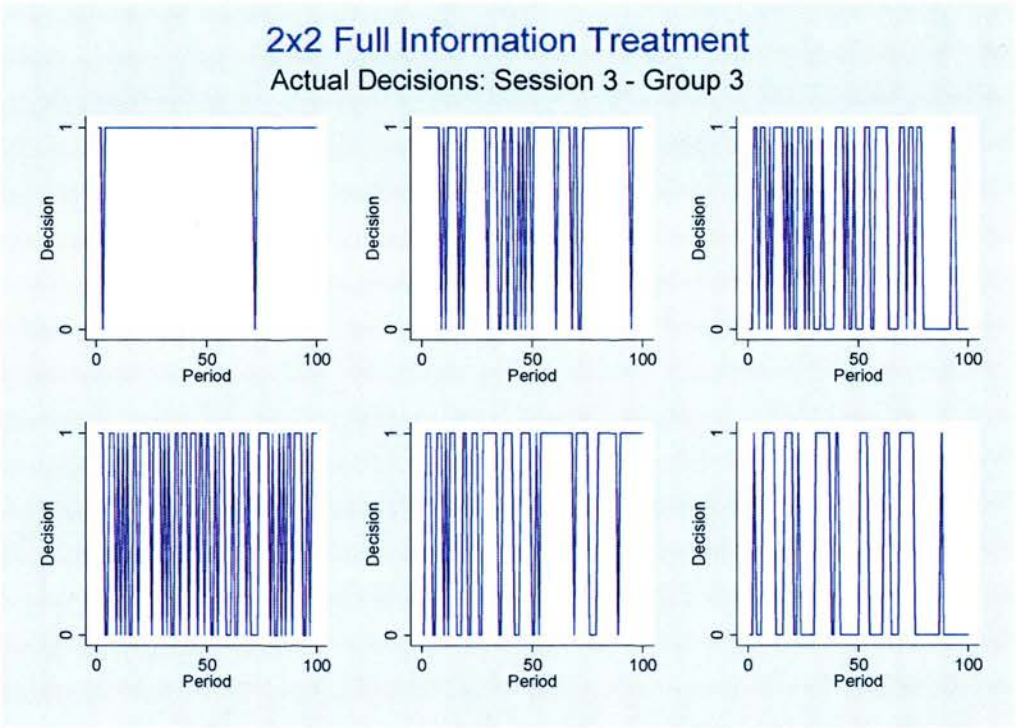

Figure 3.50: 2 × 2 Full Information Treatment (Session 3, Group 3): Individual Decisions

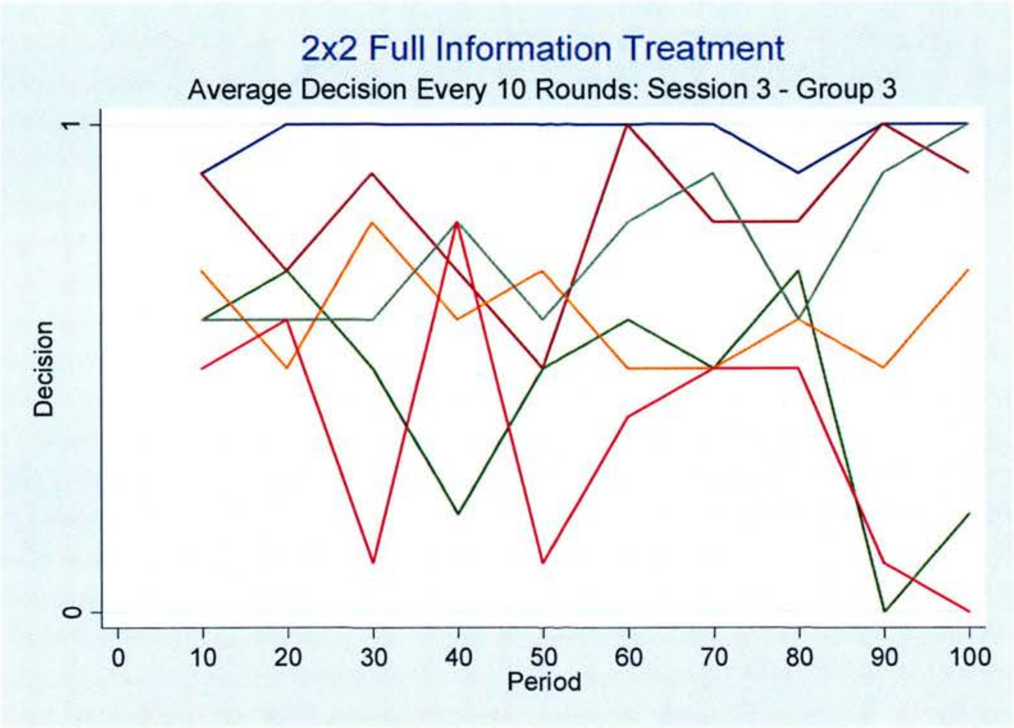

Figure 3.51: 2 × 2 Full Information Treatment (Session 3, Group 3): Average Individual Decision Every 10 Rounds

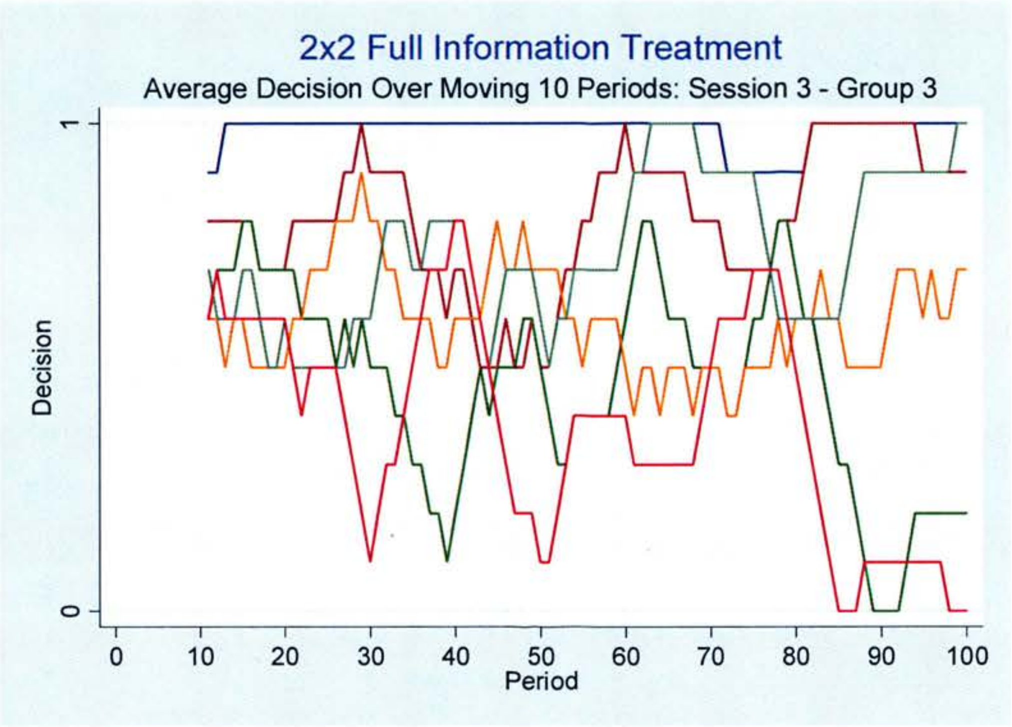

Figure 3.52: 2 × 2 Full Information Treatment (Session 3, Group 3): Average Individual Decision Over Moving 10 Rounds

Session 8, Group 1

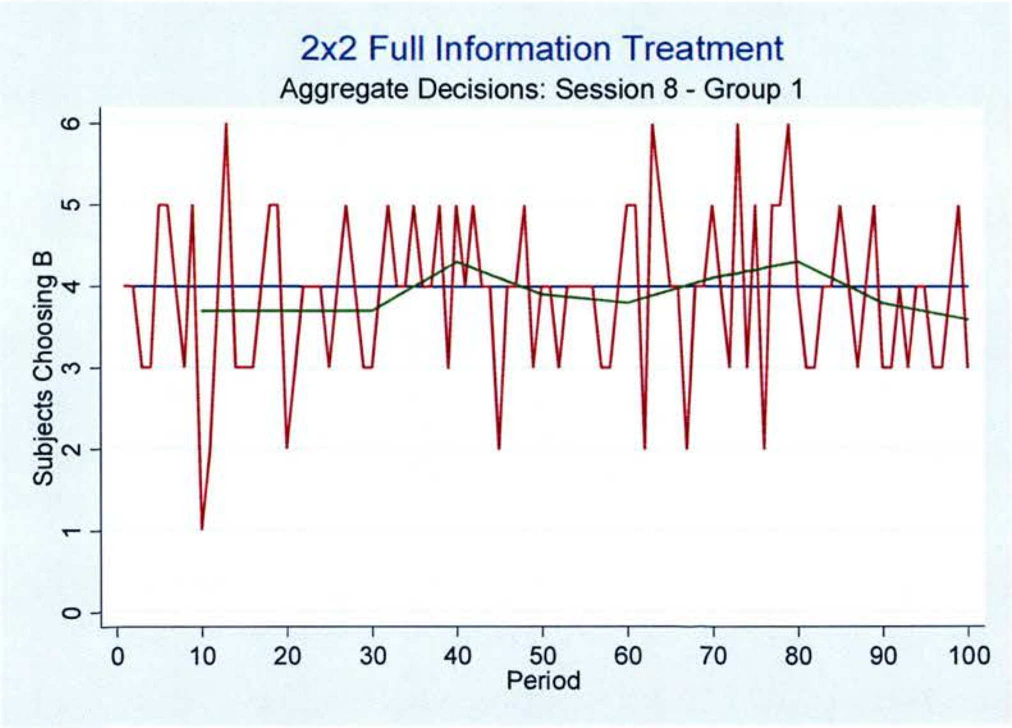

Figure 3.53: 2 × 2 Full Information Treatment (Session 8, Group 1): Aggregate Decisions

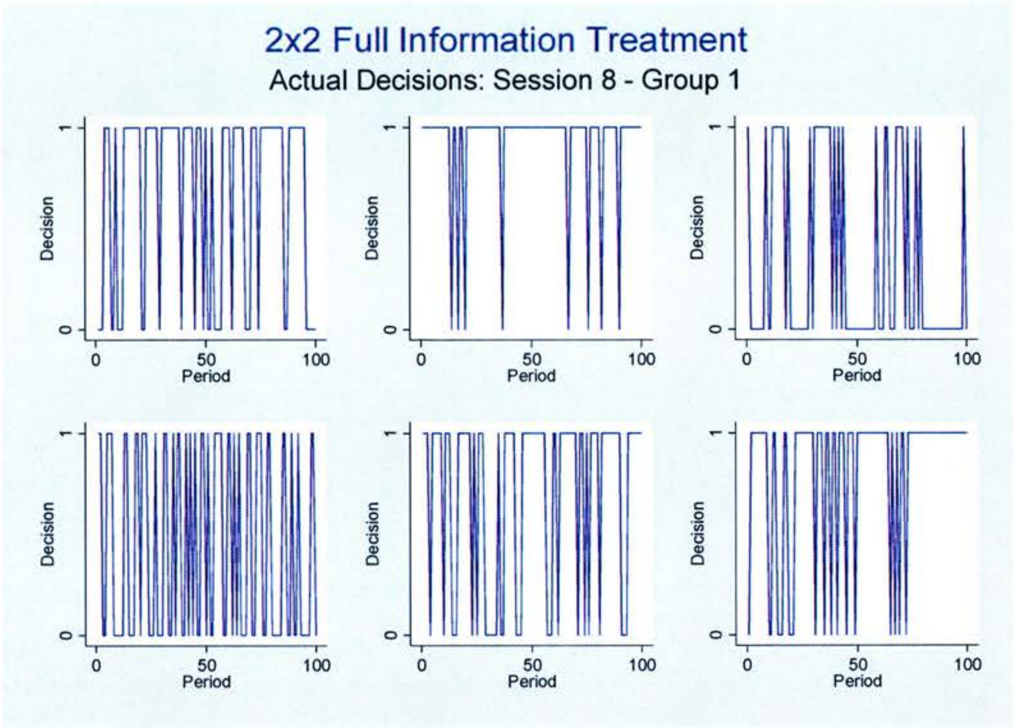

Figure 3.54: 2 × 2 Full Information Treatment (Session 8, Group 1): Individual Decisions

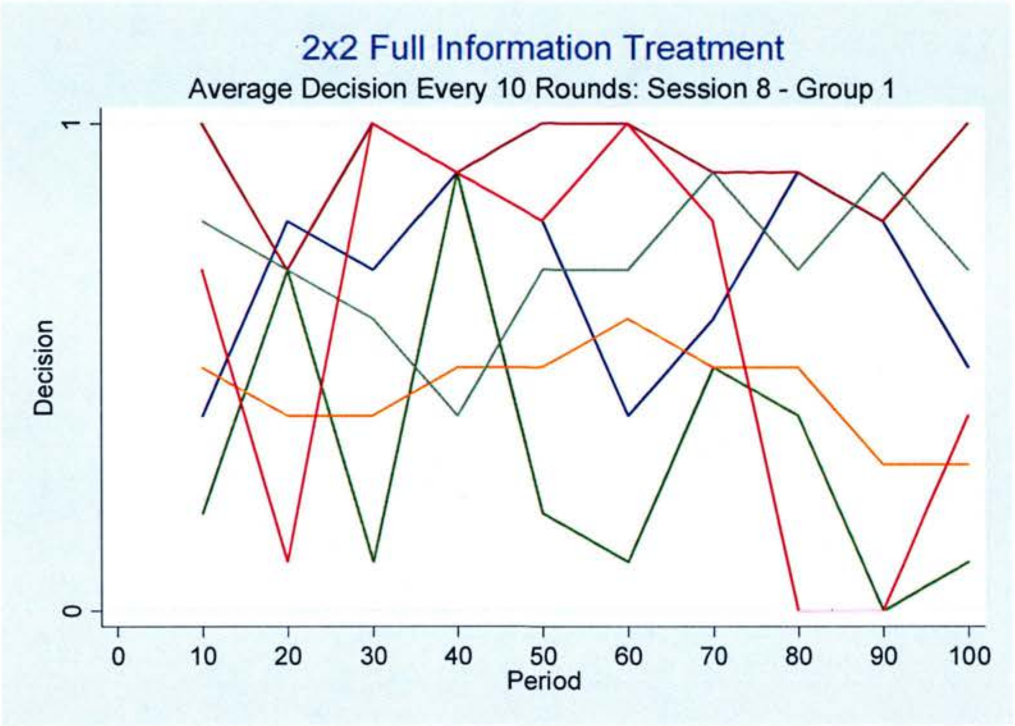

Figure 3.55: 2 × 2 Full Information Treatment (Session 8, Group 1): Average Individual Decision Every 10 Rounds

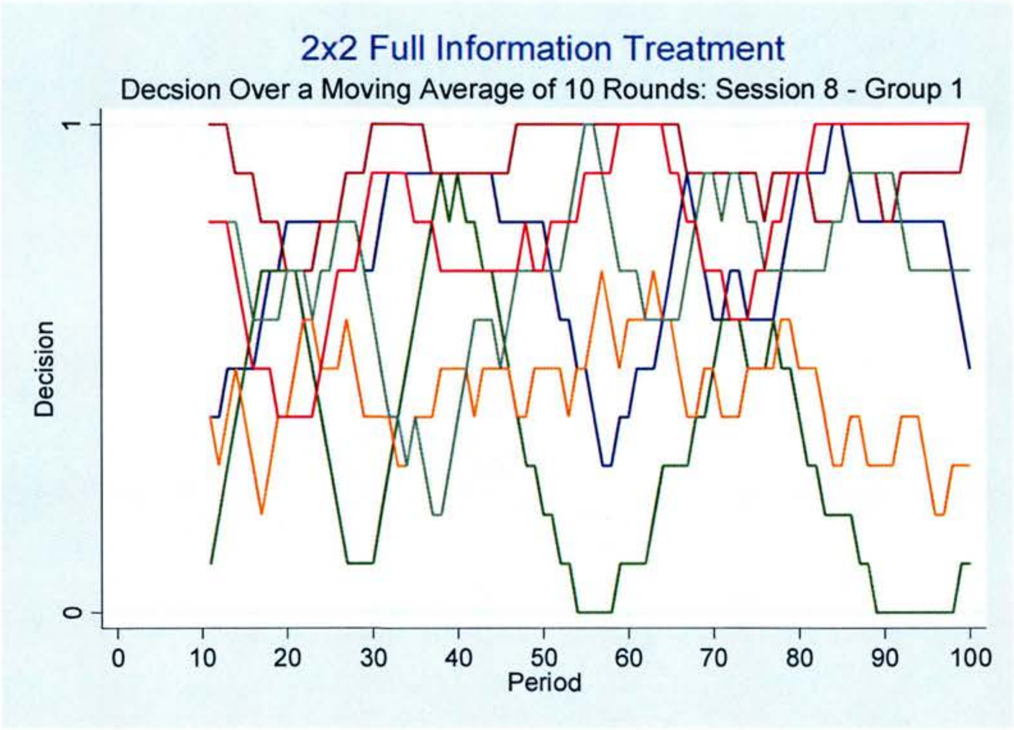

Figure 3.56: 2 × 2 Full Information Treatment (Session 8, Group 1): Average Individual Decision Over Moving 10 Rounds

Session 8, Group 2

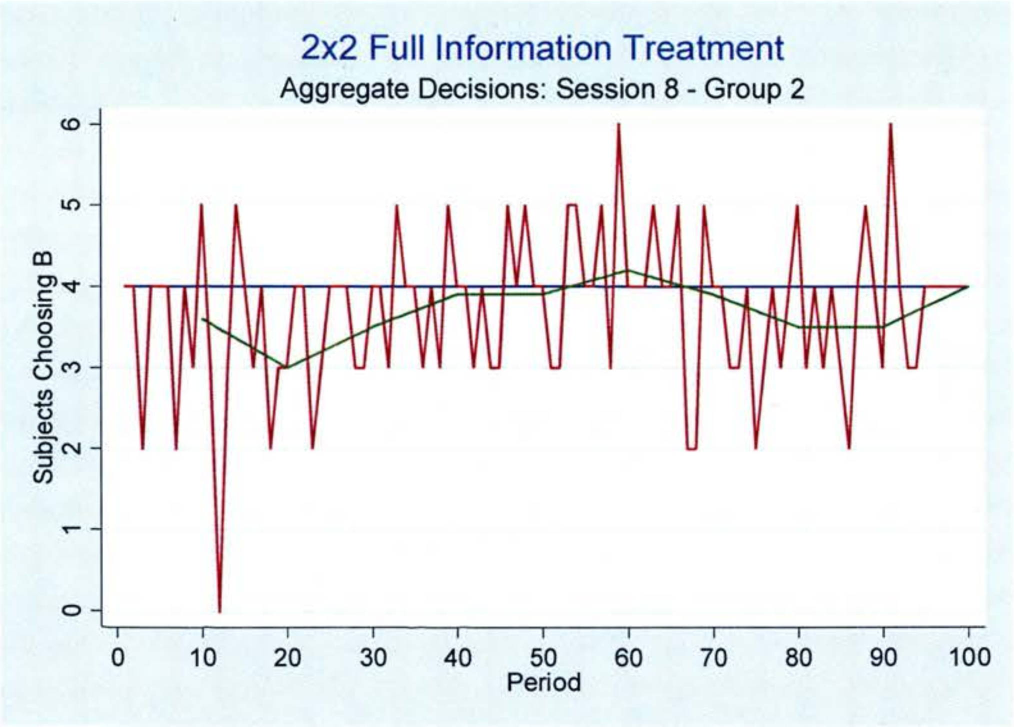

Figure 3.57: 2 × 2 Full Information Treatment (Session 8, Group 2): Aggregate Decisions

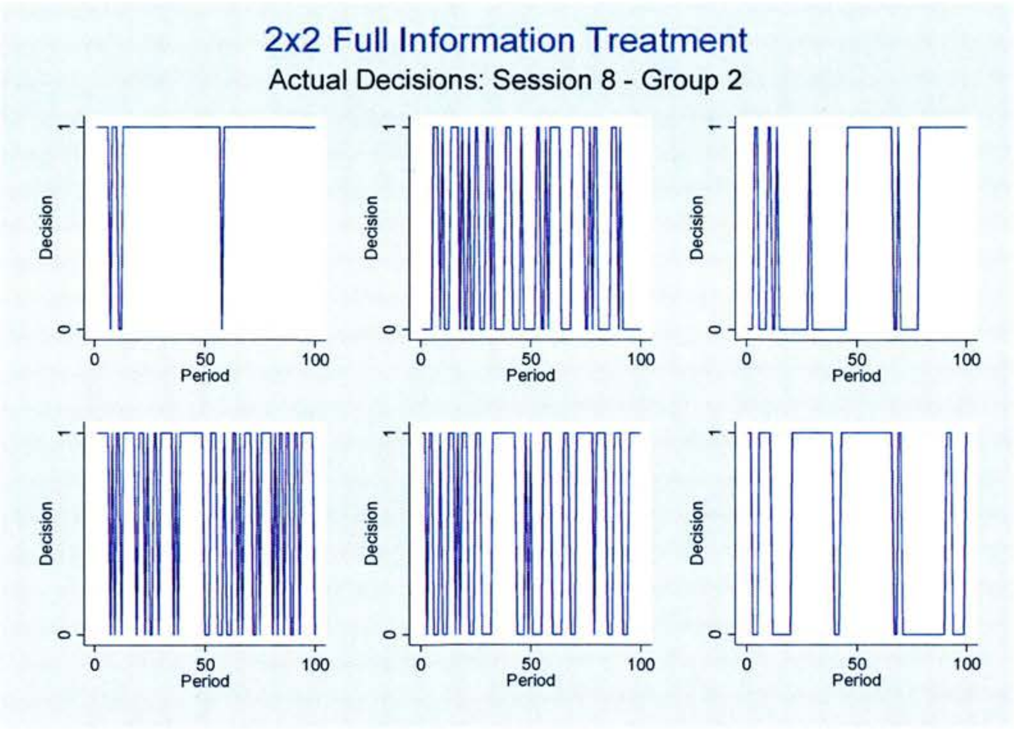

Figure 3.58: 2 × 2 Full Information Treatment (Session 8, Group 2): Individual Decisions

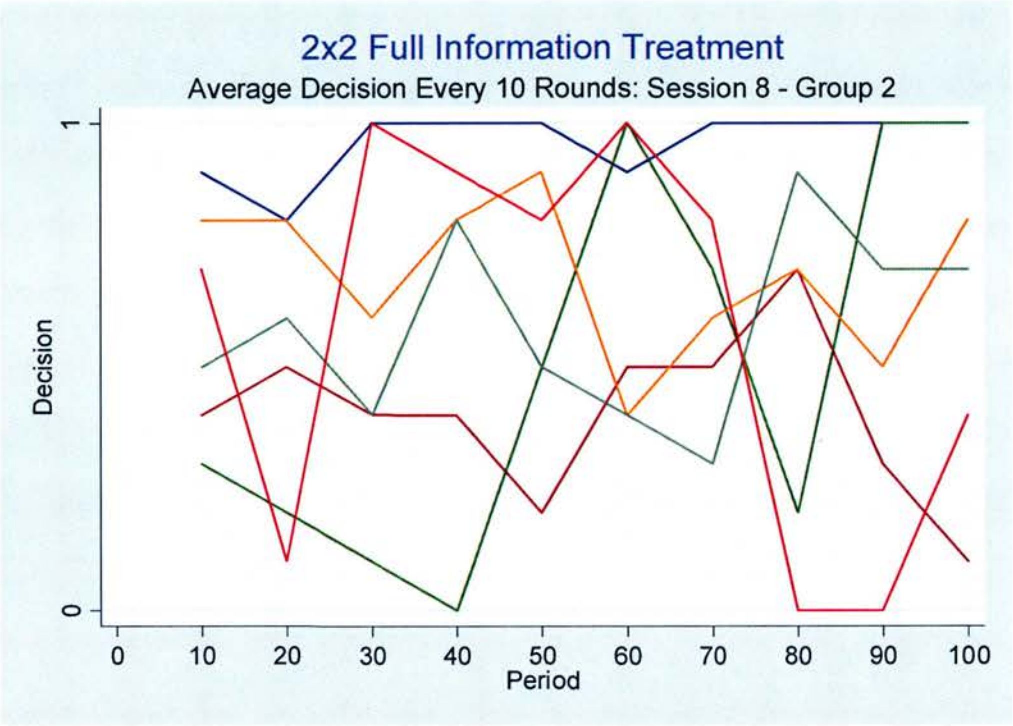

Figure 3.59: 2 × 2 Full Information Treatment (Session 8, Group 2): Average Individual Decision Every 10 Rounds

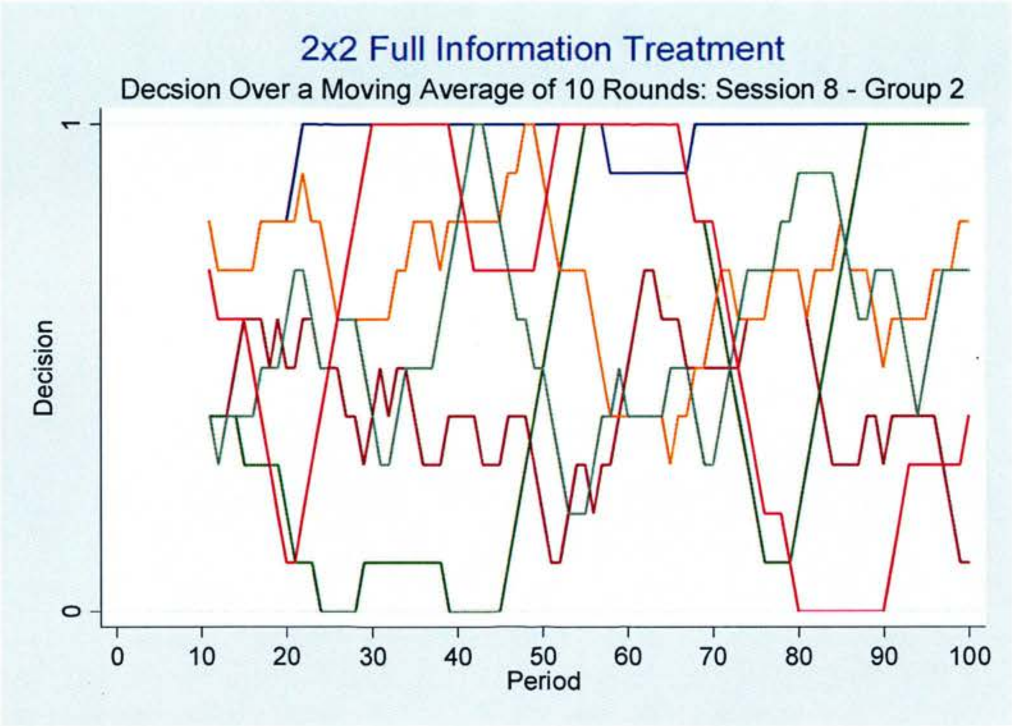

Figure 3.60: 2 × 2 Full Information Treatment (Session 8, Group 2): Average Individual Decision Over Moving 10 Rounds

3.A.3 3×3 Aggregate Information Treatment Data

Session 2, Group 1

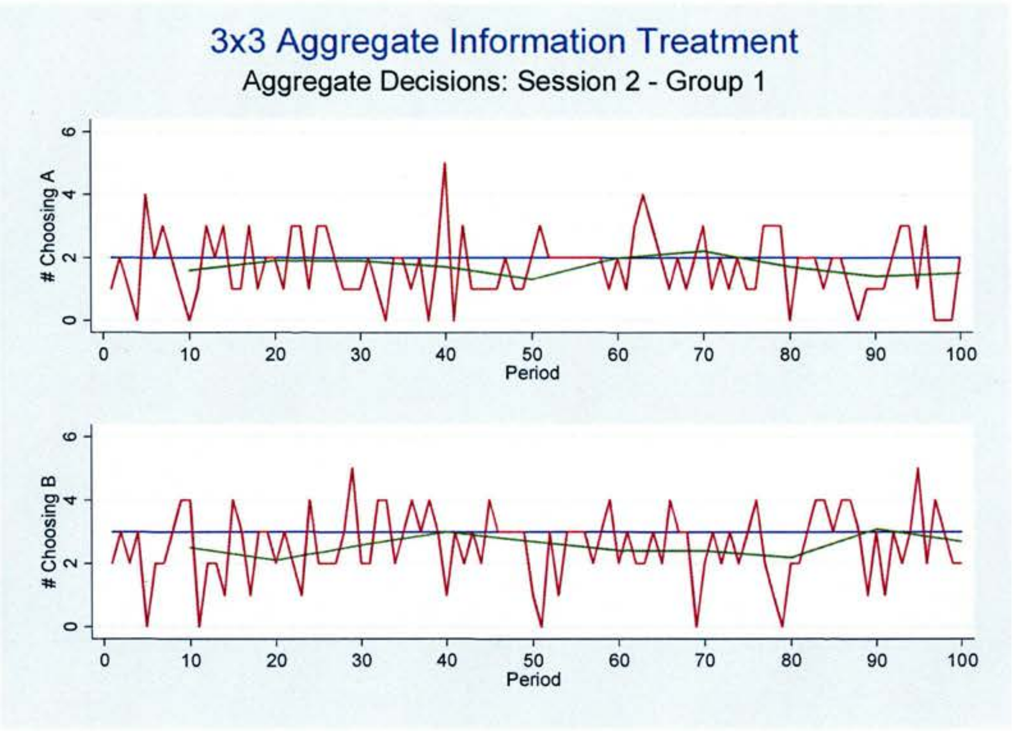

Figure 3.61: 3×3 Aggregate Information Treatment (Session 2, Group 1): Aggregate Decisions

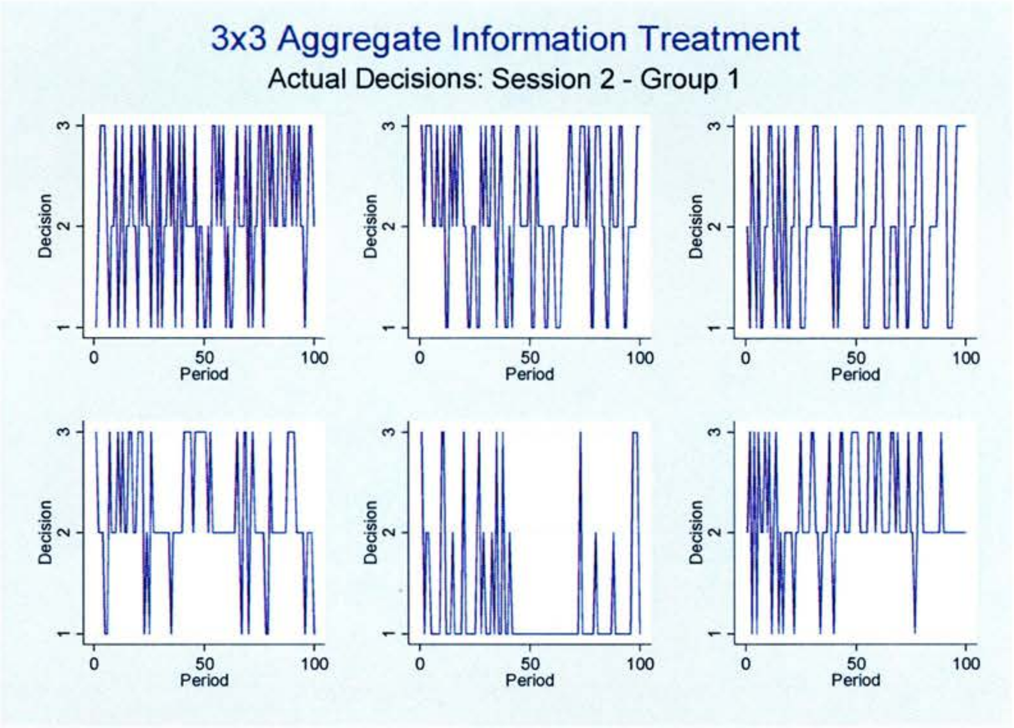

Figure 3.62: 3×3 Aggregate Information Treatment (Session 2, Group 1): Individual Decisions

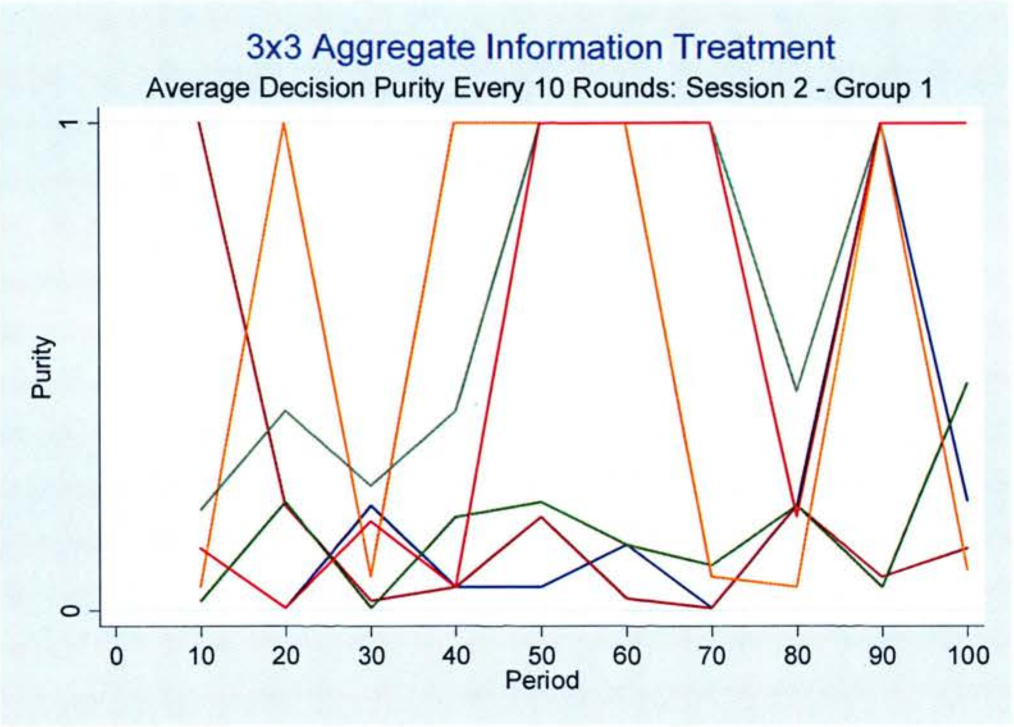

Figure 3.63: 3 × 3 Aggregate Information Treatment (Session 2, Group 1): Average Individual Decision Purity Every 10 Rounds

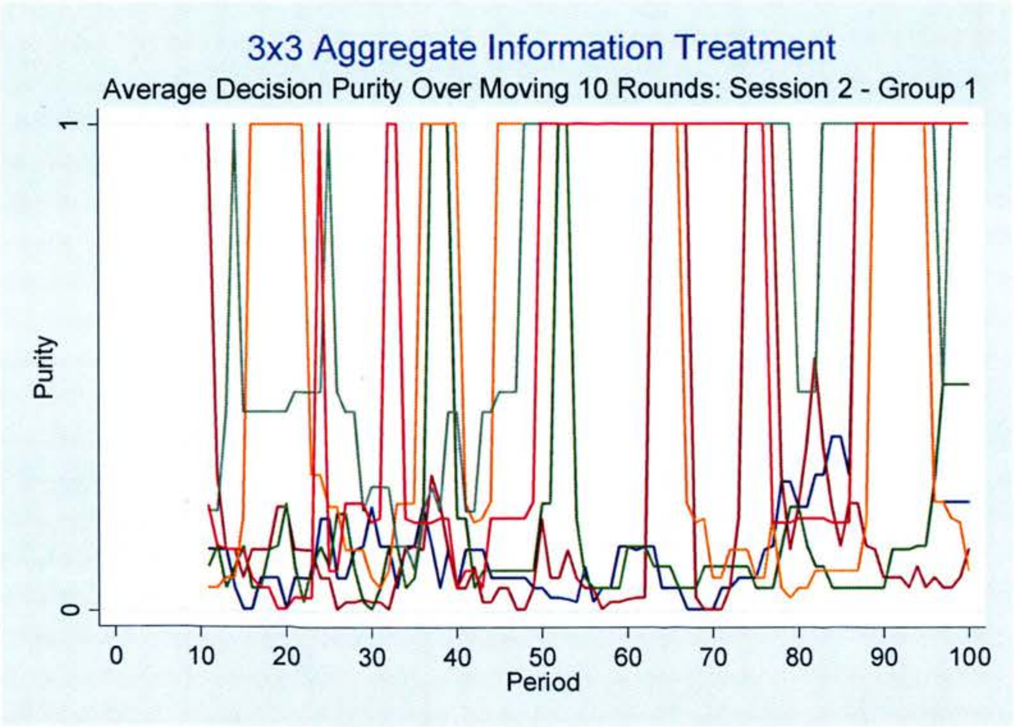

Figure 3.64: 3 × 3 Aggregate Information Treatment (Session 2, Group 1): Average Individual Decision Purity Over Moving 10 Rounds

Session 2, Group 2

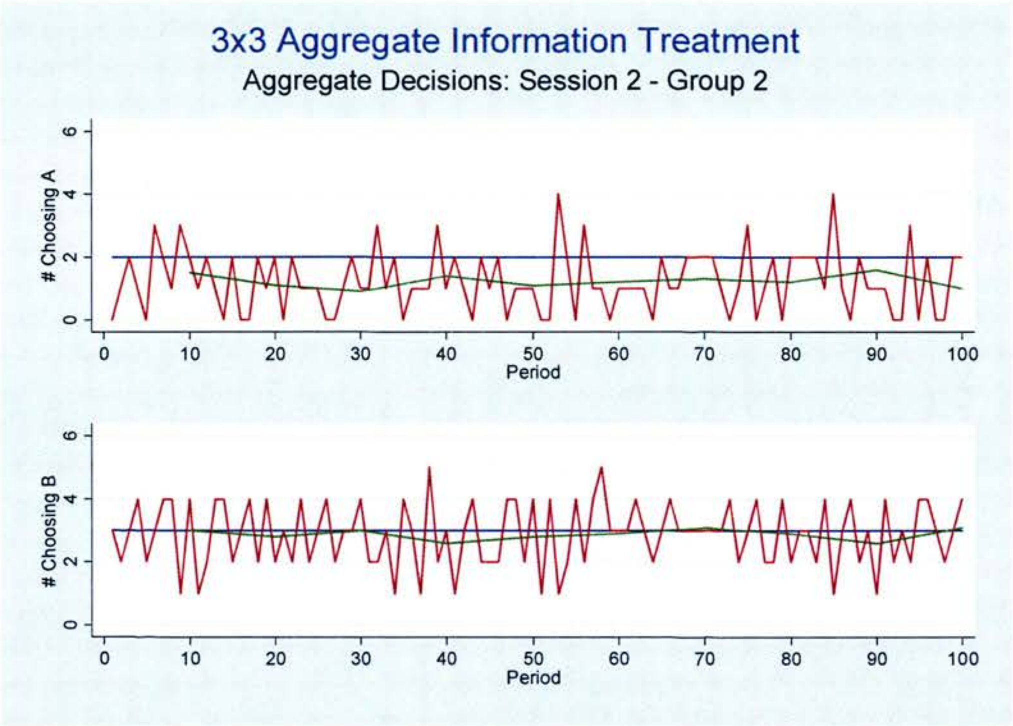

Figure 3.65: 3×3 Aggregate Information Treatment (Session 2, Group 2): Aggregate Decisions

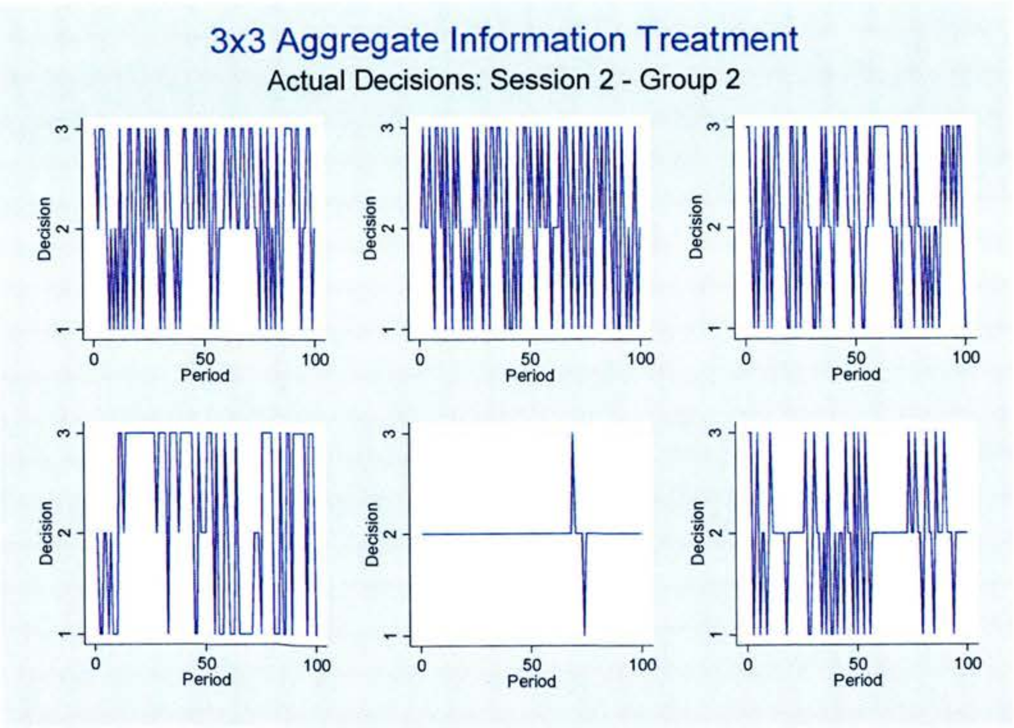

Figure 3.66: 3×3 Aggregate Information Treatment (Session 2, Group 2): Individual Decisions

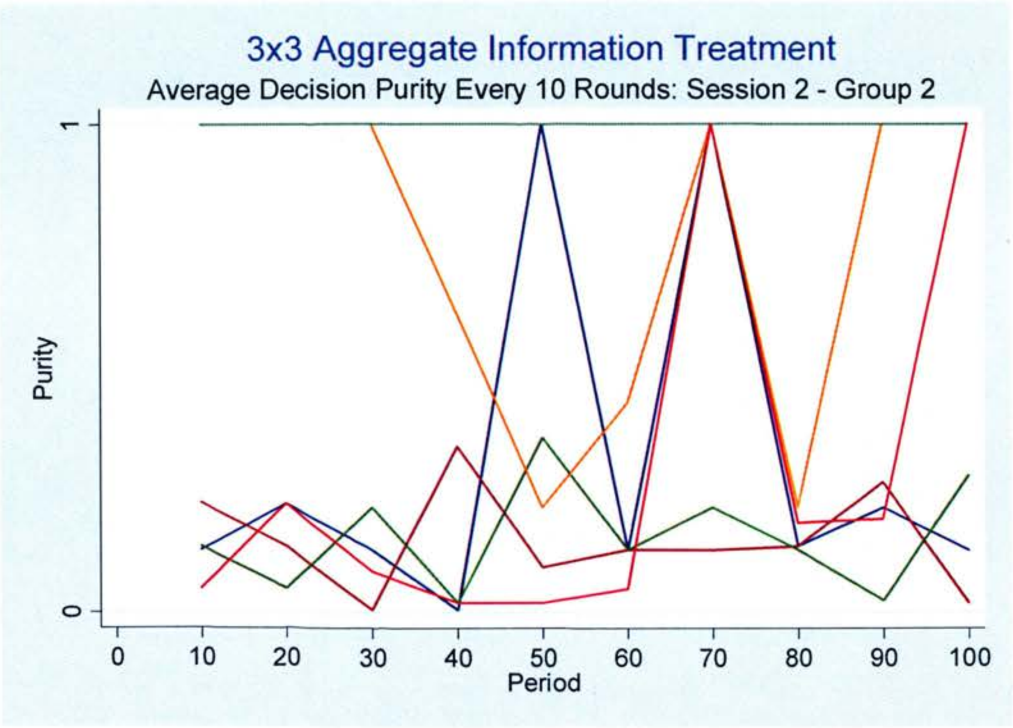

Figure 3.67: 3 × 3 Aggregate Information Treatment (Session 2, Group 2): Average Individual Decision Purity Every 10 Rounds

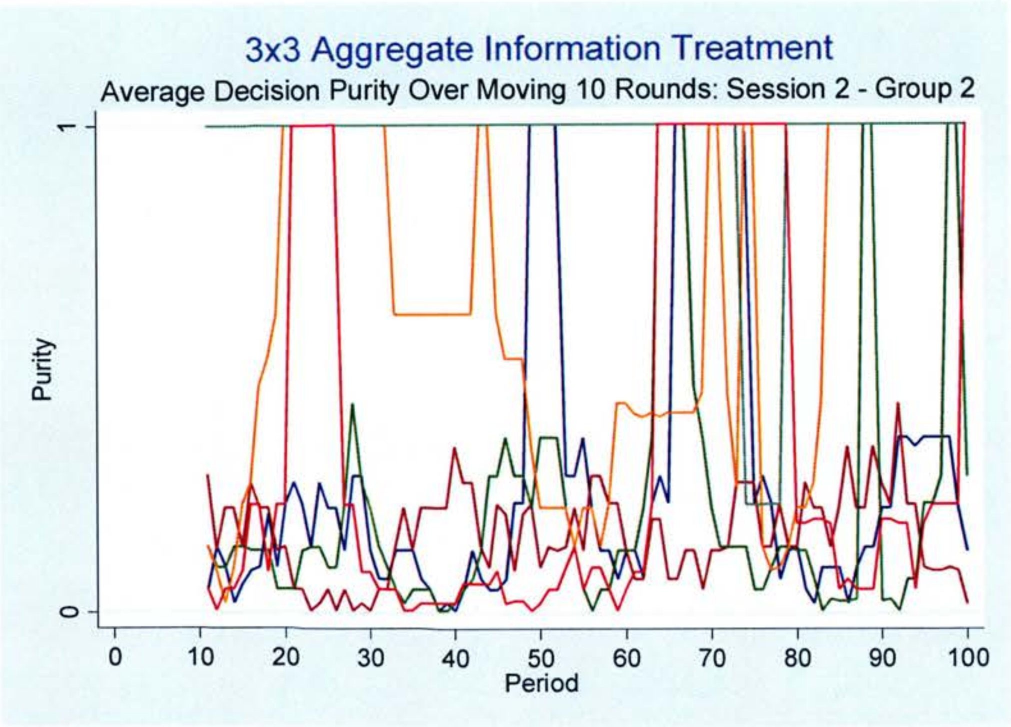

Figure 3.68: 3 × 3 Aggregate Information Treatment (Session 2, Group 2): Average Individual Decision Purity Over Moving 10 Rounds

Session 2, Group 3

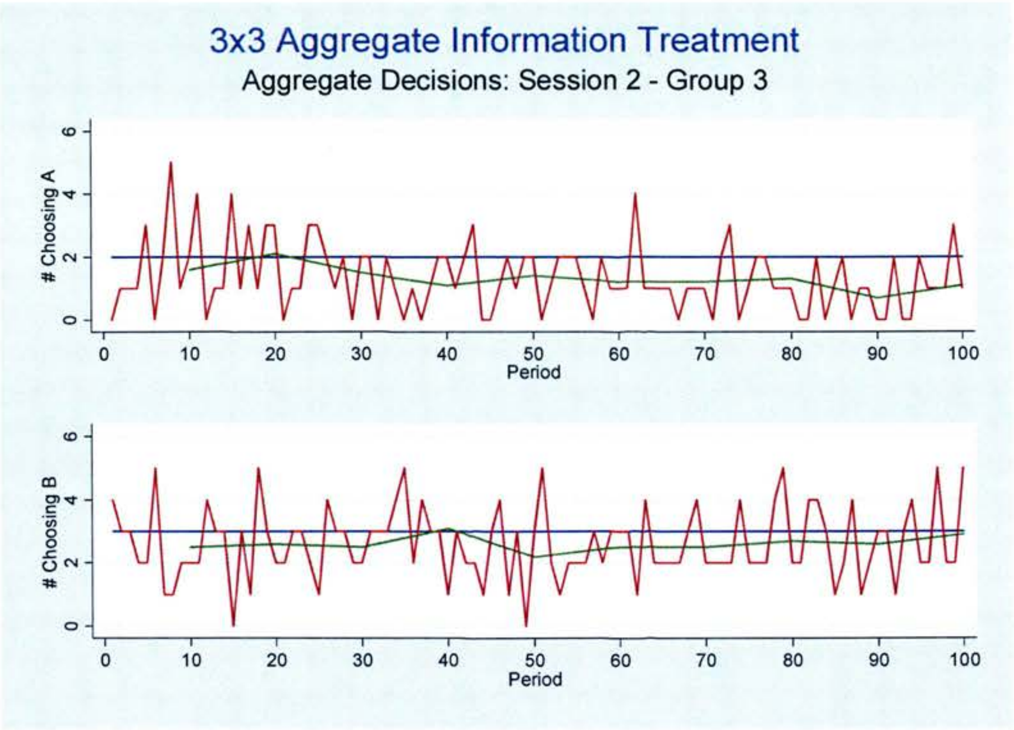

Figure 3.69: 3×3 Aggregate Information Treatment (Session 2, Group 3): Aggregate Decisions

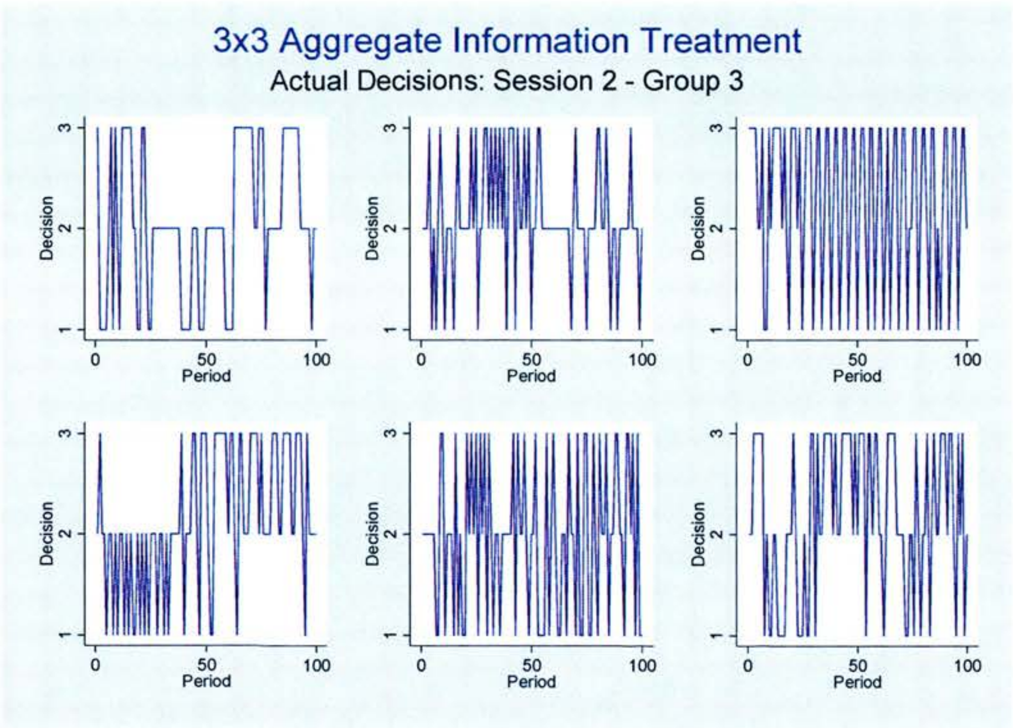

Figure 3.70: 3 × 3 Aggregate Information Treatment (Session 2, Group 3): Individual Decisions

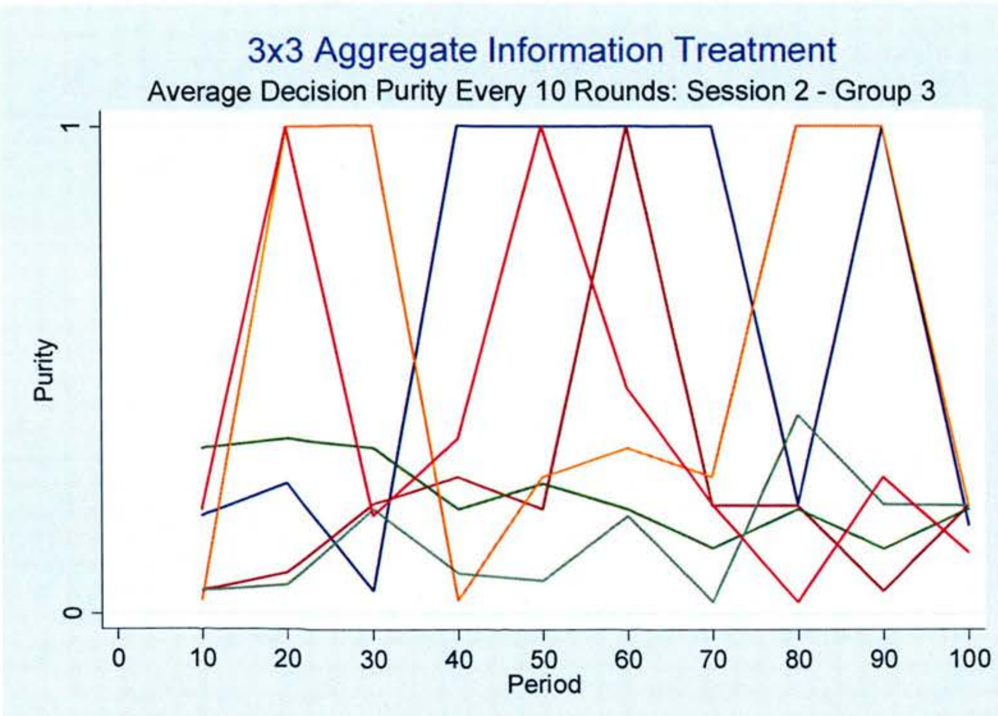

Figure 3.71: 3 × 3 Aggregate Information Treatment (Session 2, Group 3): Average Individual Decision Purity Every 10 Rounds

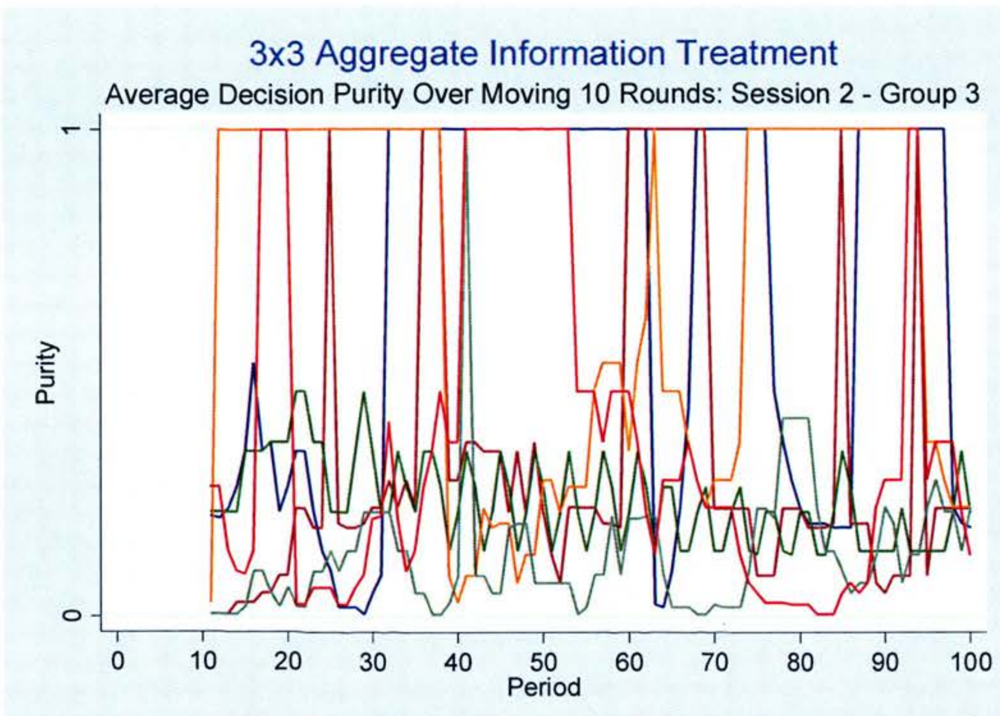

Figure 3.72: 3 × 3 Aggregate Information Treatment (Session 2, Group 3): Average Individual Decision Purity Over Moving 10 Rounds

Session 6, Group 1

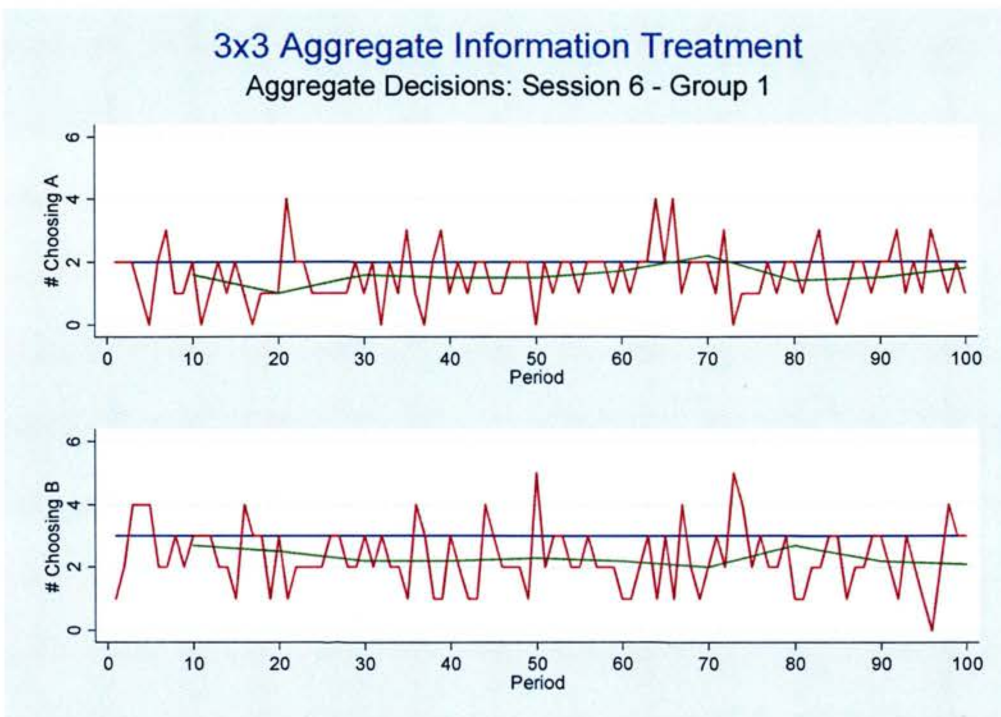

Figure 3.73: 3×3 Aggregate Information Treatment (Session 6, Group 1): Aggregate Decisions

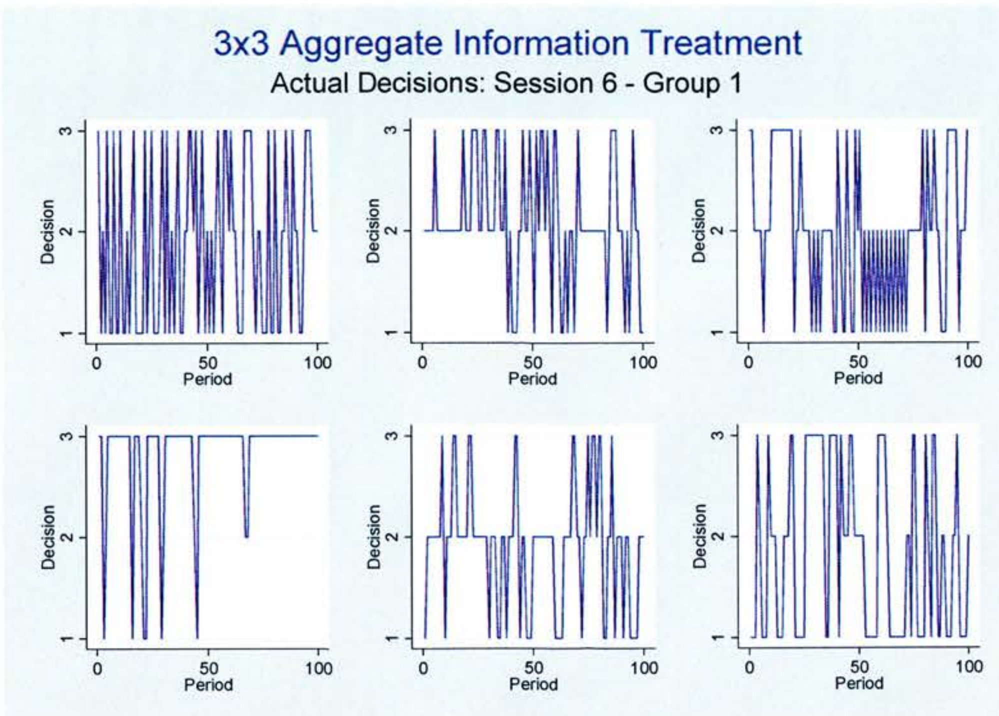

Figure 3.74: 3×3 Aggregate Information Treatment (Session 6, Group 1): Individual Decisions

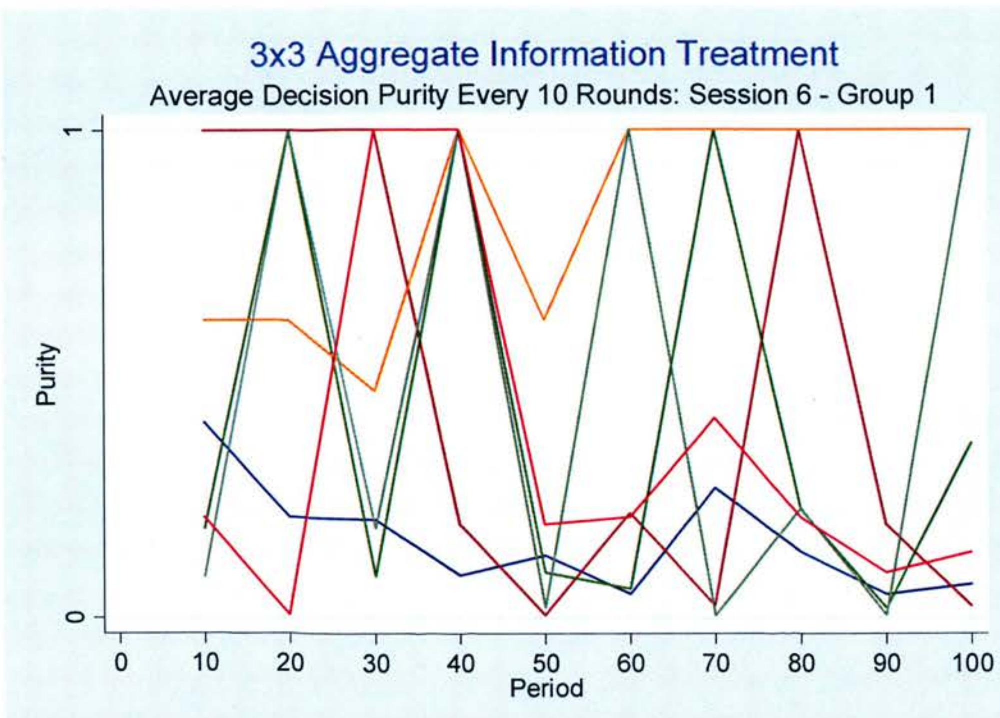

Figure 3.75: 3×3 Aggregate Information Treatment (Session 6, Group 1): Average Individual Decision Purity Every 10 Rounds

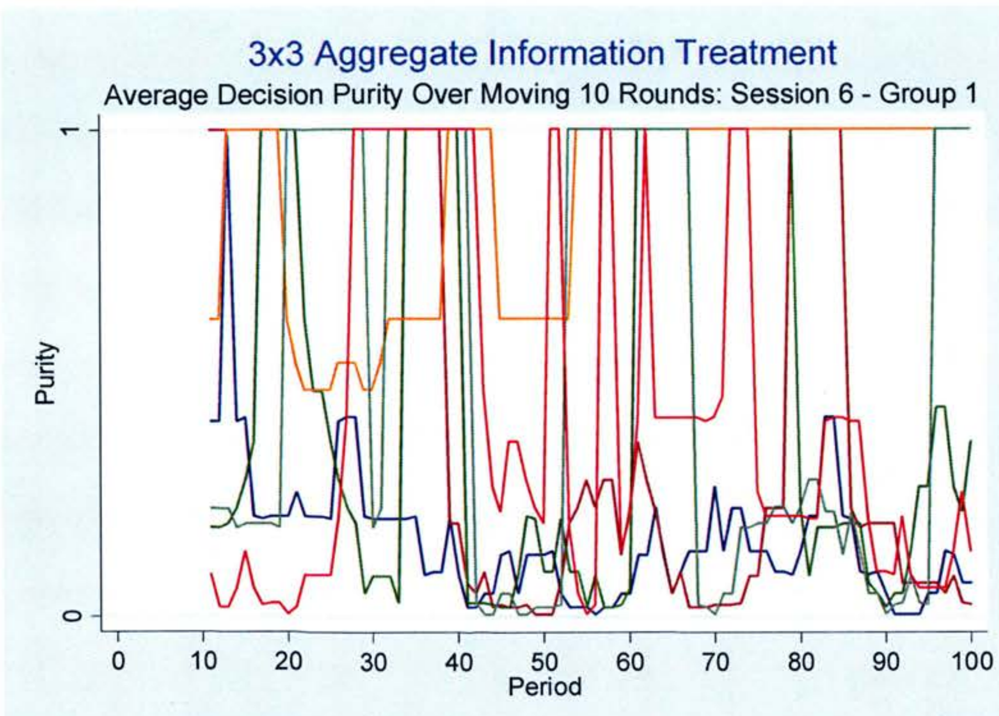

Figure 3.76: 3×3 Aggregate Information Treatment (Session 6, Group 1): Average Individual Decision Purity Over Moving 10 Rounds

Session 6, Group 2

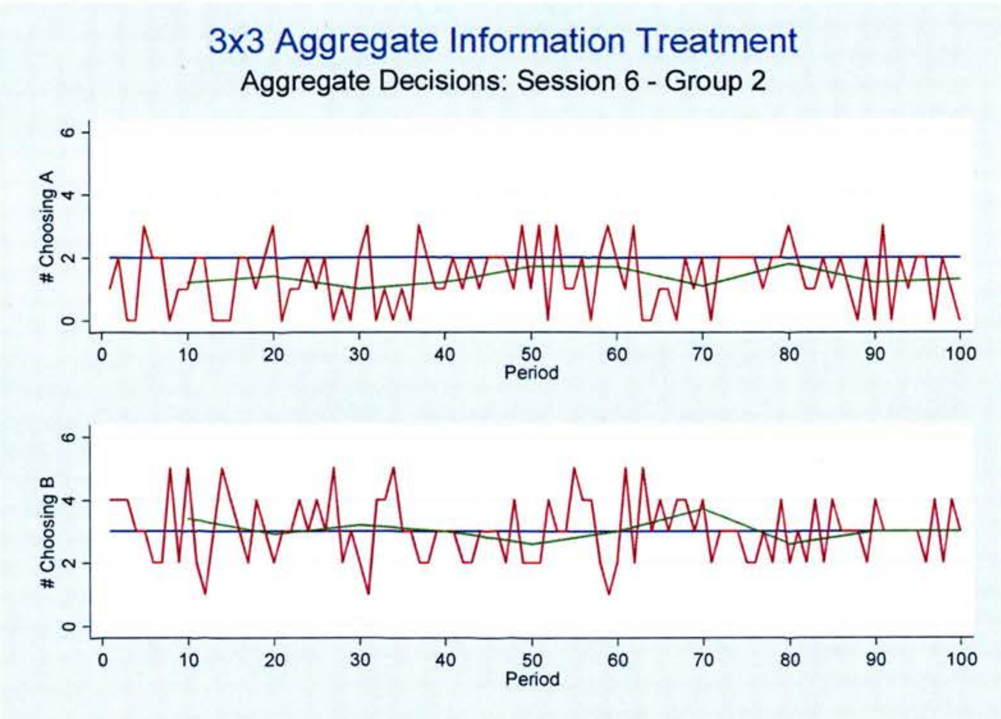

Figure 3.77: 3×3 Aggregate Information Treatment (Session 6, Group 2): Aggregate Decisions

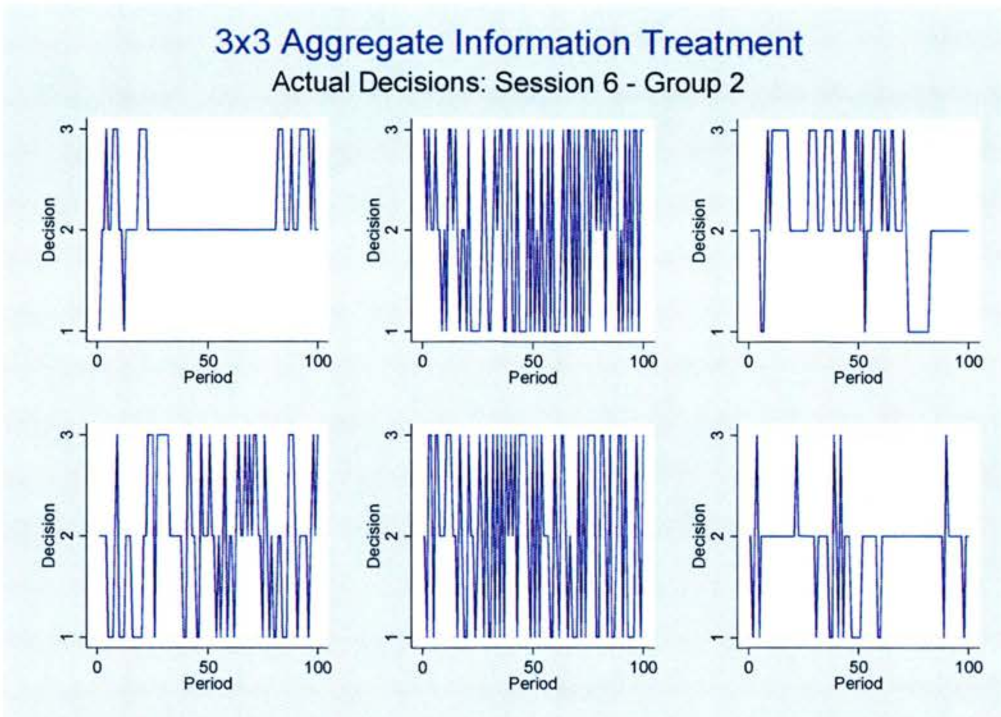

Figure 3.78: 3 × 3 Aggregate Information Treatment (Session 6, Group 2): Individual Decisions

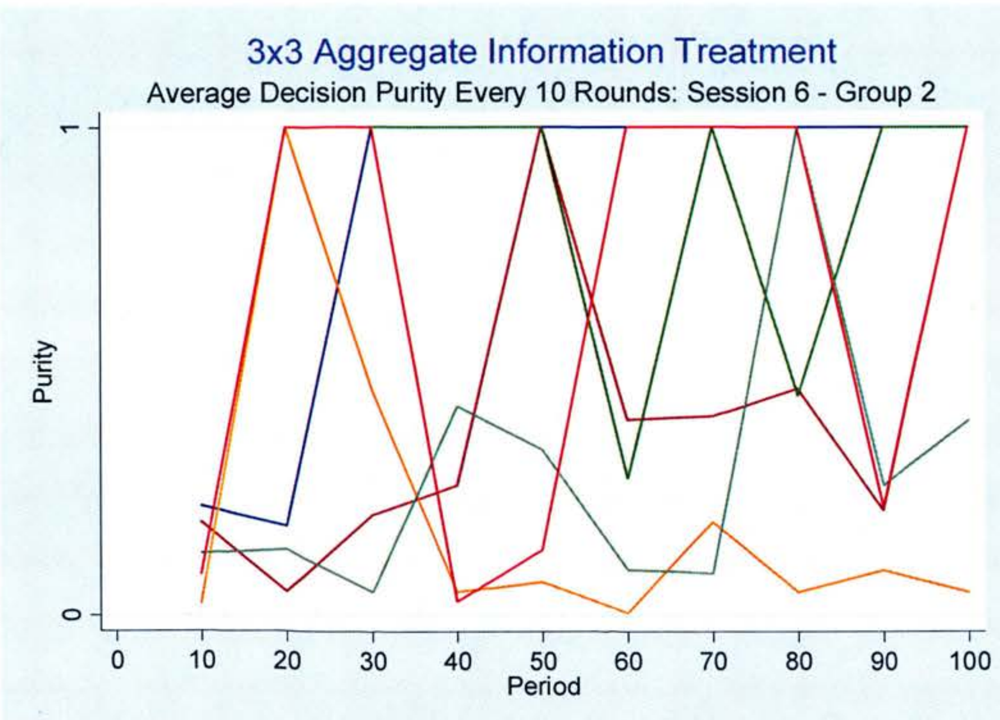

Figure 3.79: 3 × 3 Aggregate Information Treatment (Session 6, Group 2): Average Individual Decision Purity Every 10 Rounds

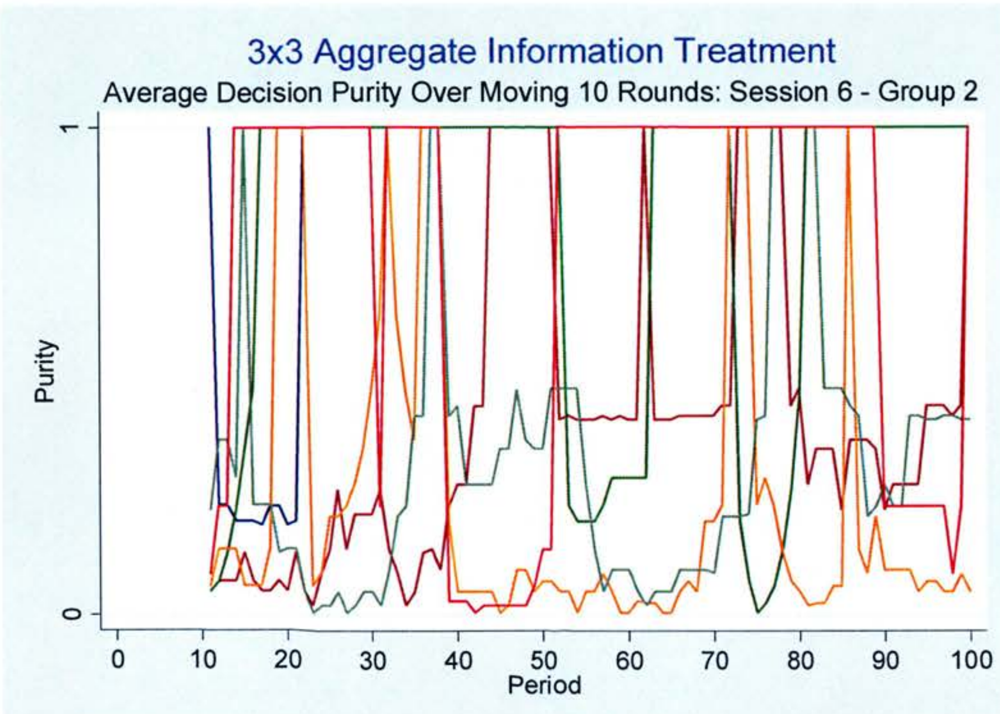

Figure 3.80: 3×3 Aggregate Information Treatment (Session 6, Group 2): Average Individual Decision Purity Over Moving 10 Rounds

Session 6, Group 3

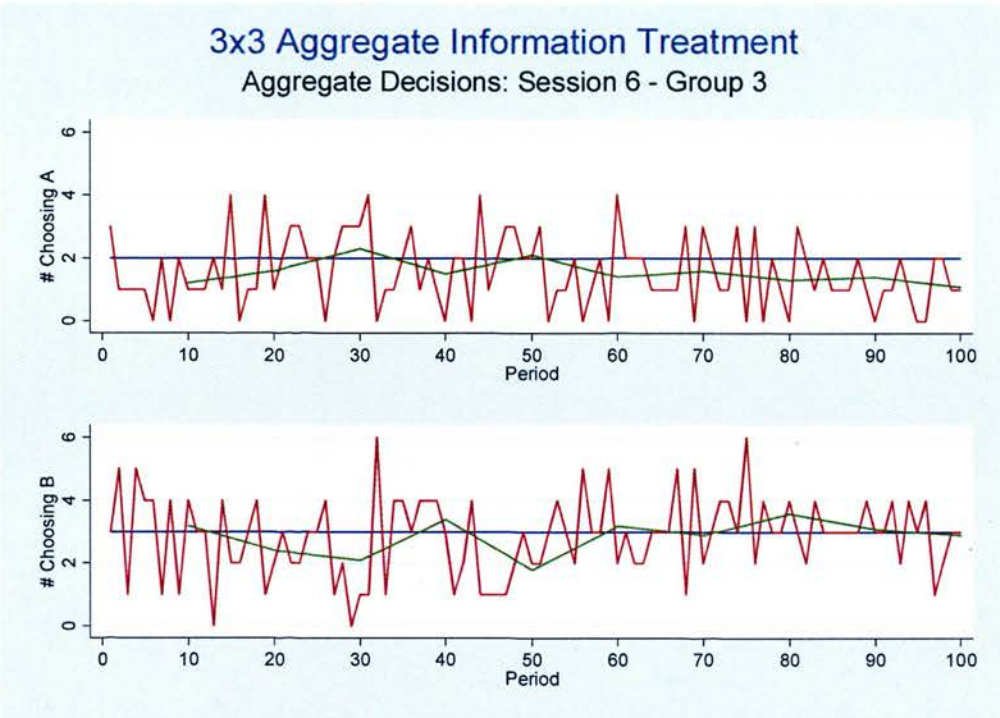

Figure 3.81: 3×3 Aggregate Information Treatment (Session 6, Group 3): Aggregate Decisions

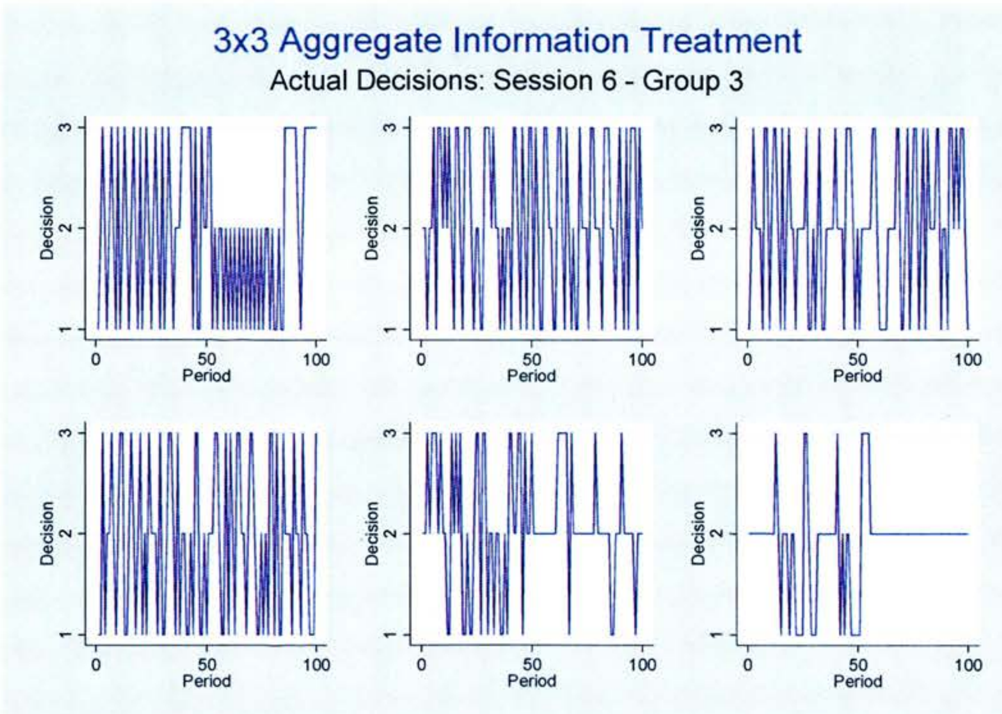

Figure 3.82: 3×3 Aggregate Information Treatment (Session 6, Group 3): Individual Decisions

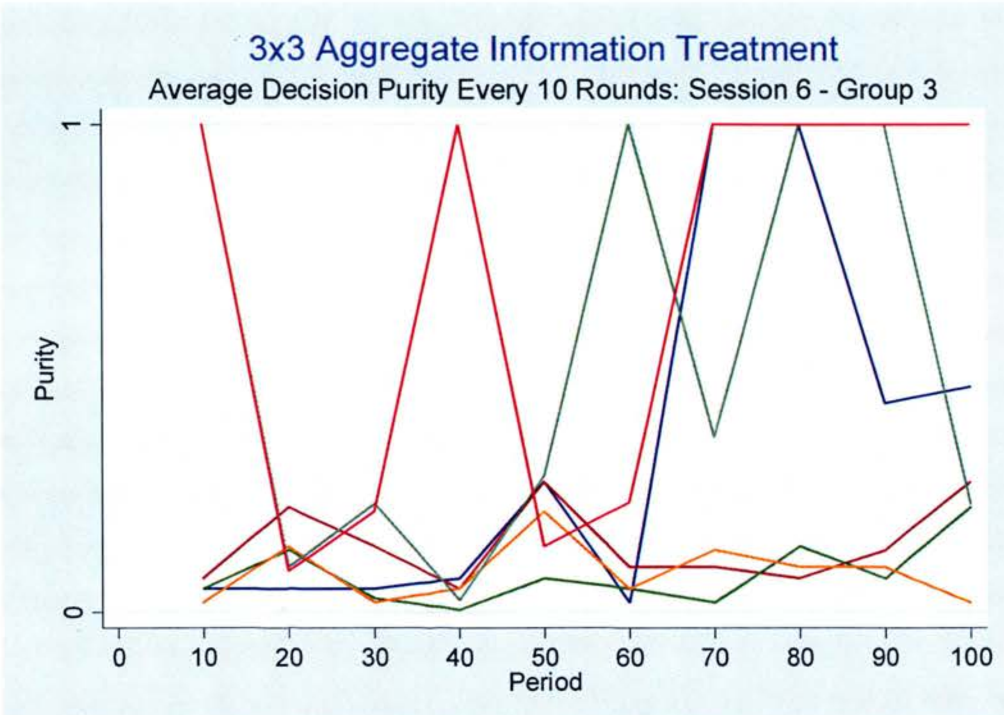

Figure 3.83: 3 × 3 Aggregate Information Treatment (Session 6, Group 3): Average Individual Decision Purity Every 10 Rounds

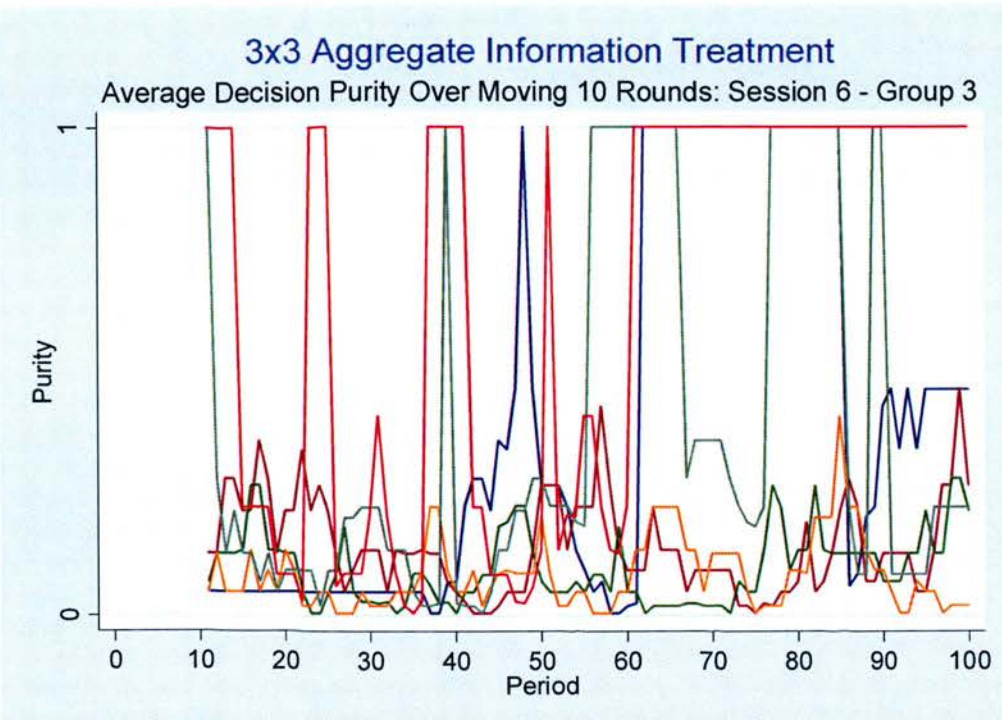

Figure 3.84: 3 × 3 Aggregate Information Treatment (Session 6, Group 3): Average Individual Decision Purity Over Moving 10 Rounds

3.A.4 3×3 Full Information Treatment Data

Session 4, Group 1

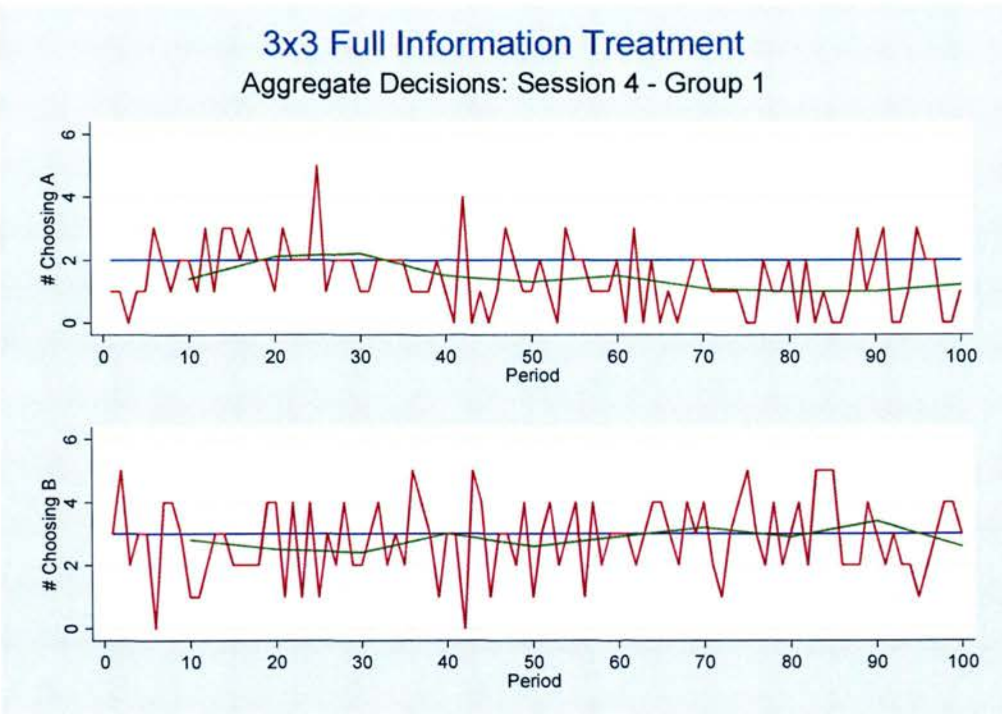

Figure 3.85: 3 × 3 Full Information Treatment (Session 4, Group 1): Aggregate Decisions

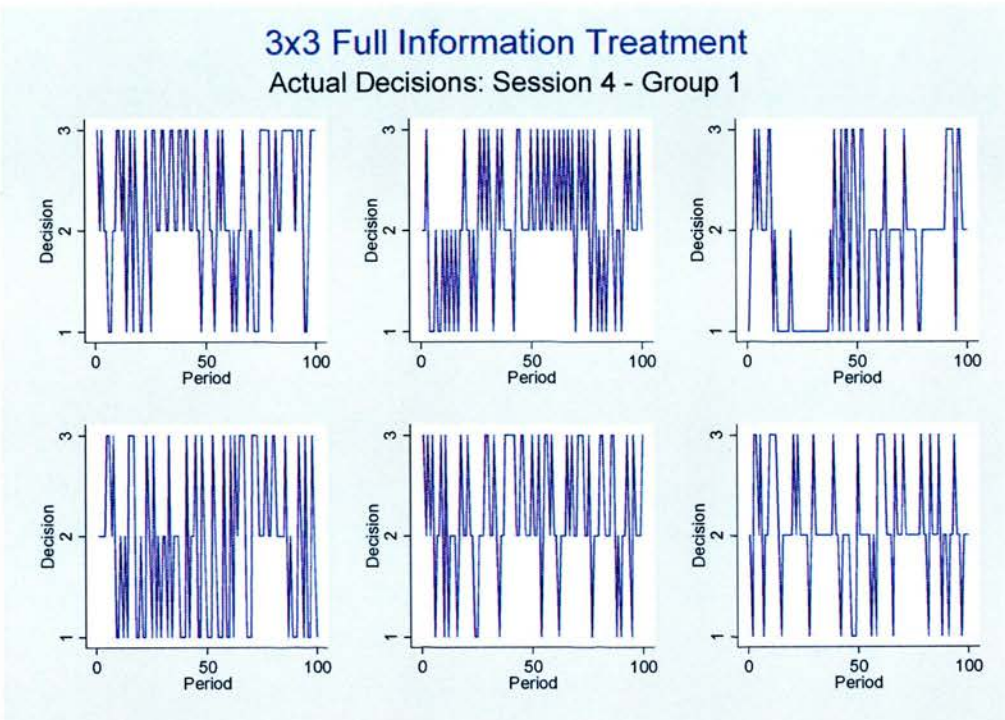

Figure 3.86: 3 × 3 Full Information Treatment (Session 4, Group 1): Individual Decisions

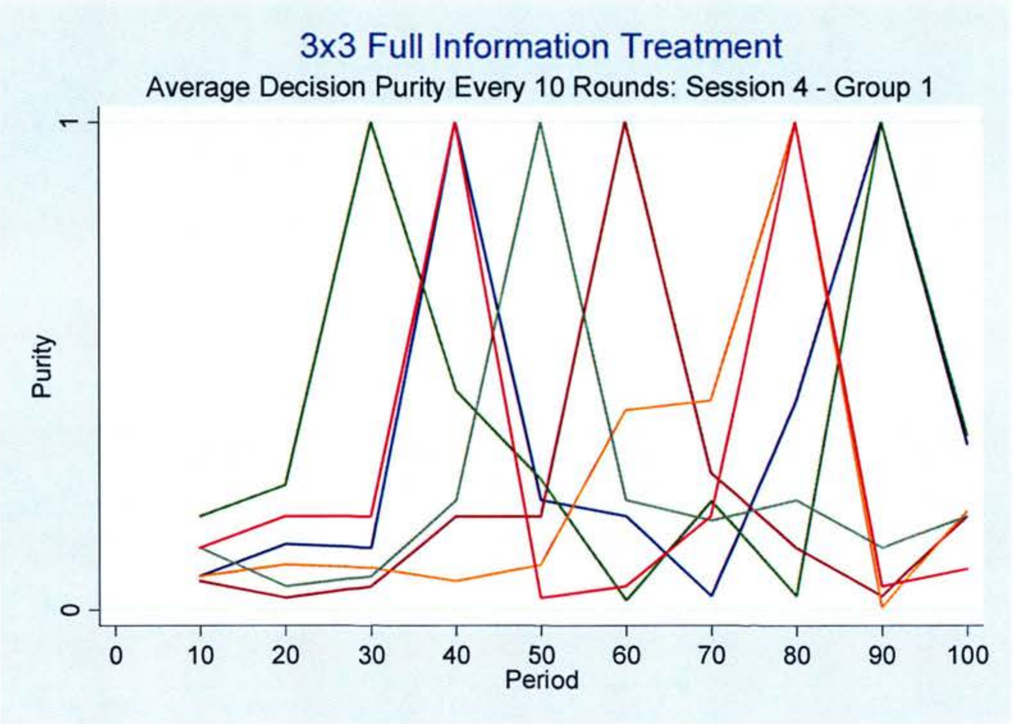

Figure 3.87: 3 × 3 Full Information Treatment (Session 4, Group 1): Average Individual Decision Purity Every 10 Rounds

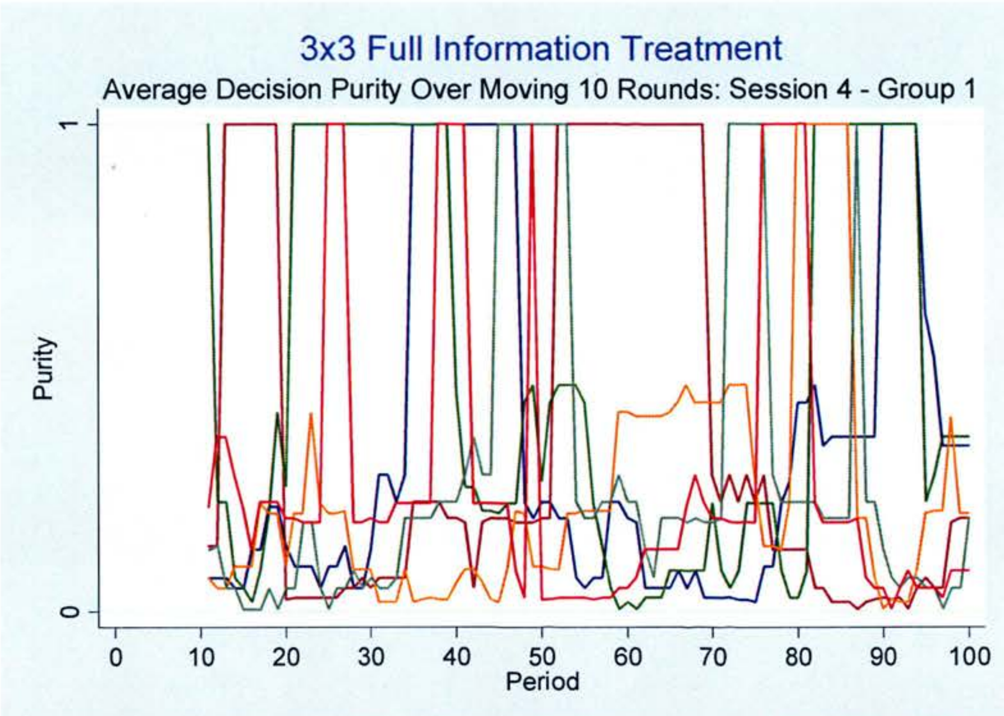

Figure 3.88: 3 × 3 Full Information Treatment (Session 4, Group 1): Average Individual Decision Purity Over Moving 10 Rounds

Session 4, Group 2

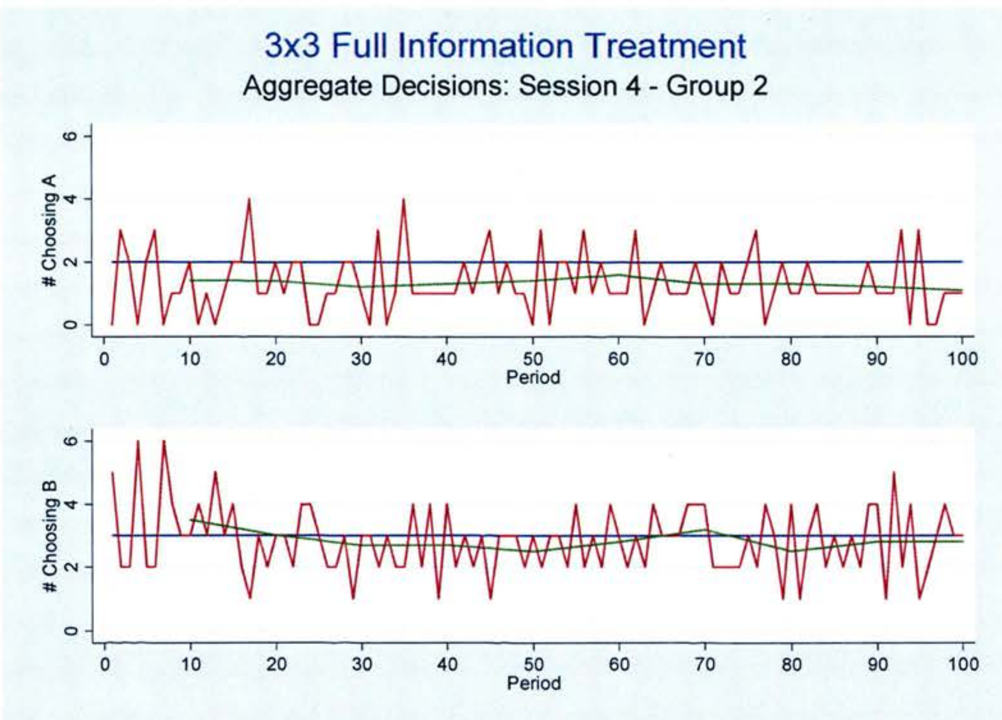

Figure 3.89: 3 × 3 Full Information Treatment (Session 4, Group 2): Aggregate Decisions

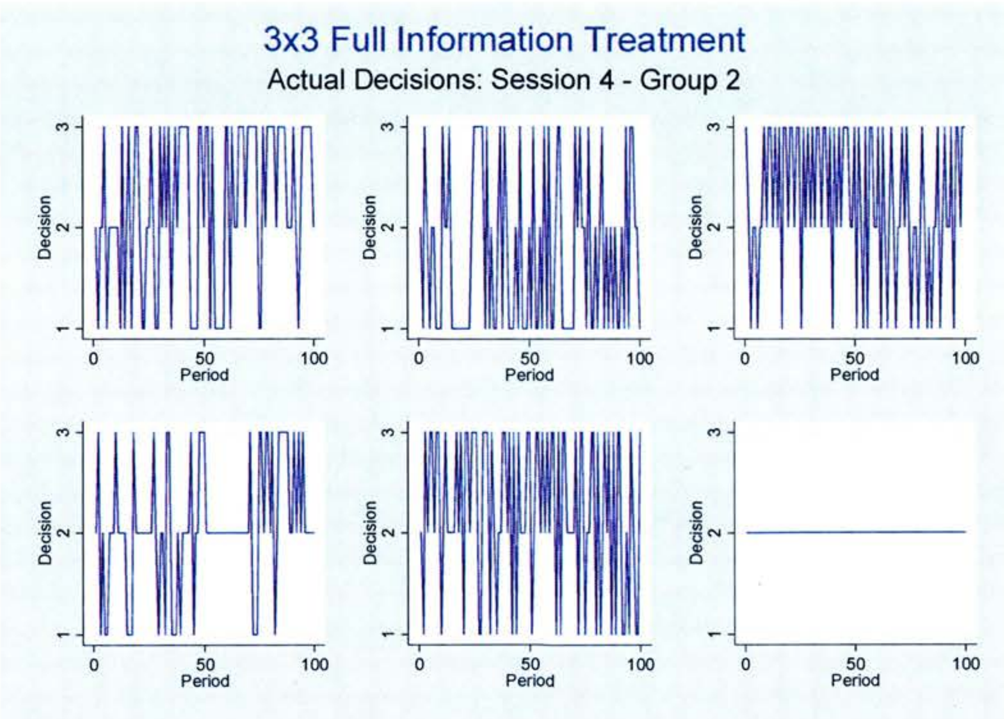

Figure 3.90: 3 × 3 Full Information Treatment (Session 4, Group 2): Individual Decisions

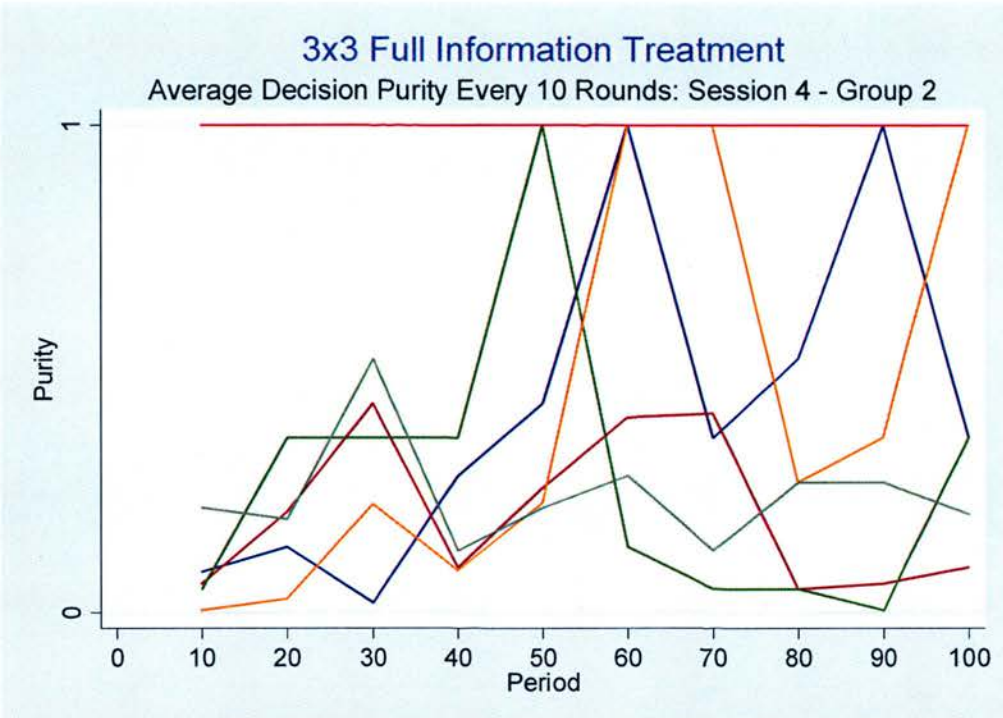

Figure 3.91: 3 × 3 Full Information Treatment (Session 4, Group 2): Average Individual Decision Purity Every 10 Rounds

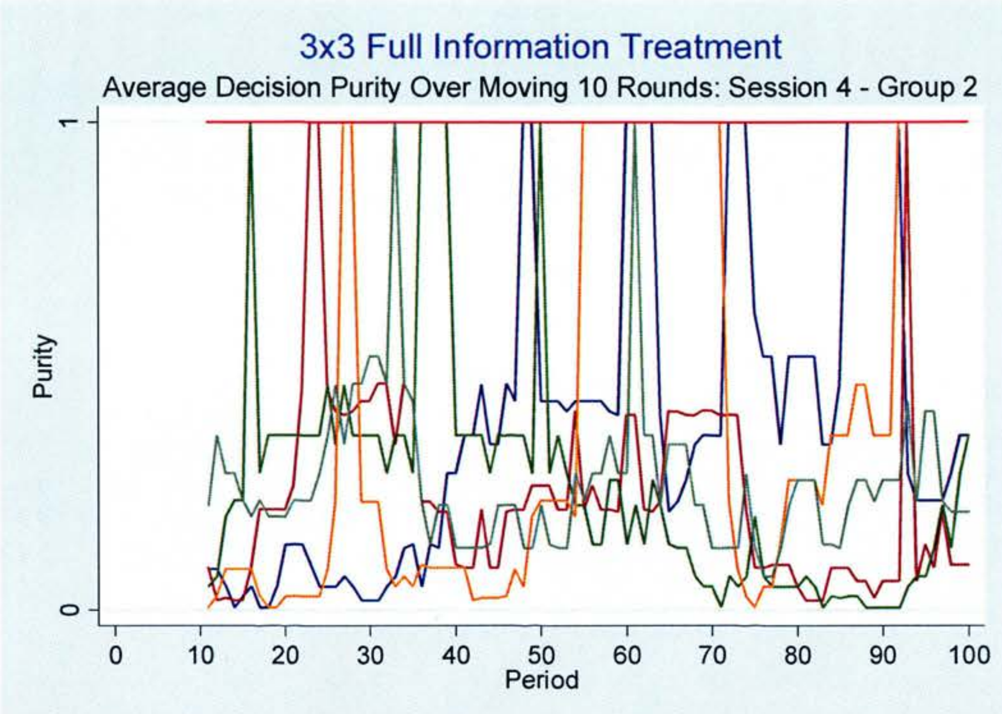

Figure 3.92: 3 × 3 Full Information Treatment (Session 4, Group 2): Average Individual Decision Purity Over Moving 10 Rounds

Session 5, Group 1

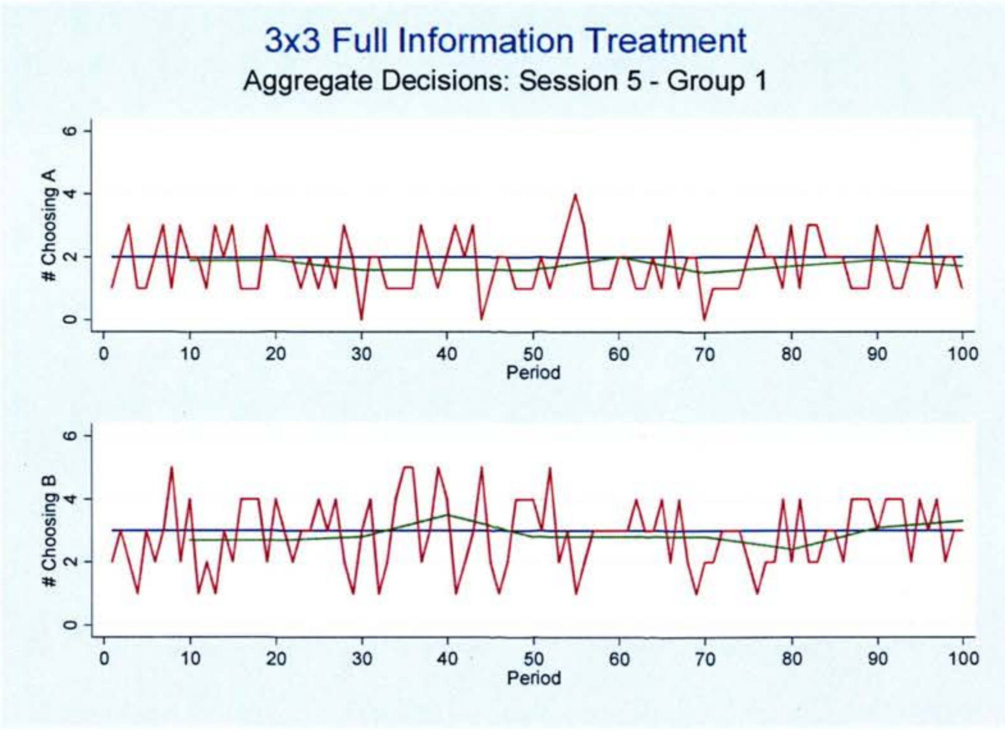

Figure 3.93: 3 × 3 Full Information Treatment (Session 5, Group 1): Aggregate Decisions

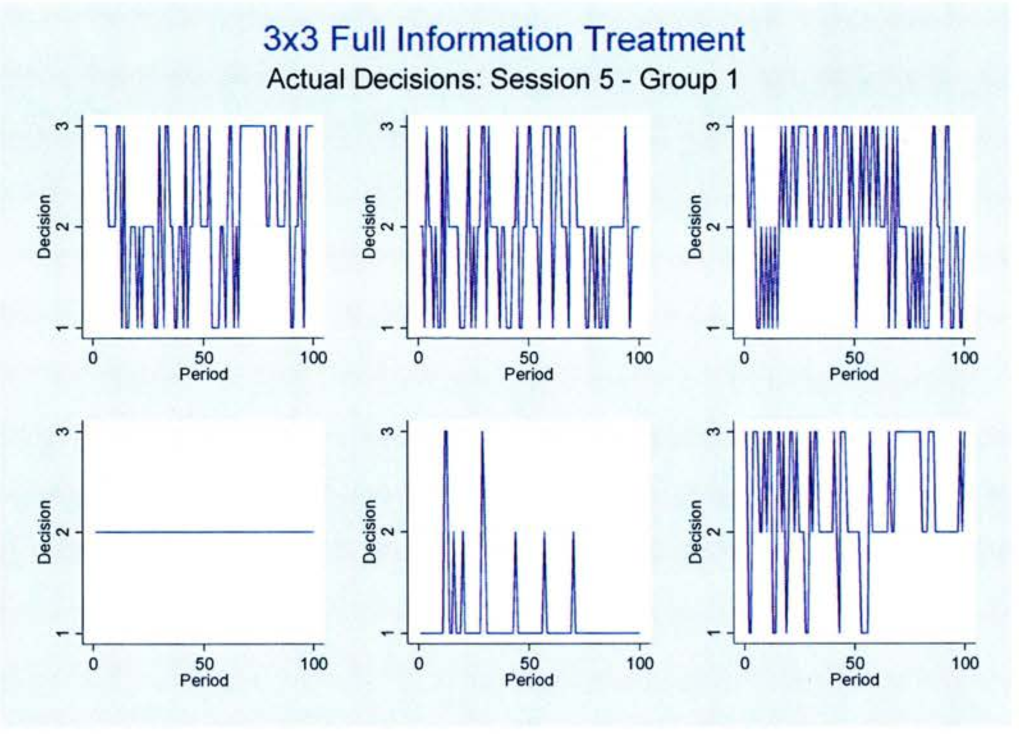

Figure 3.94: 3 × 3 Full Information Treatment (Session 5, Group 1): Individual Decisions

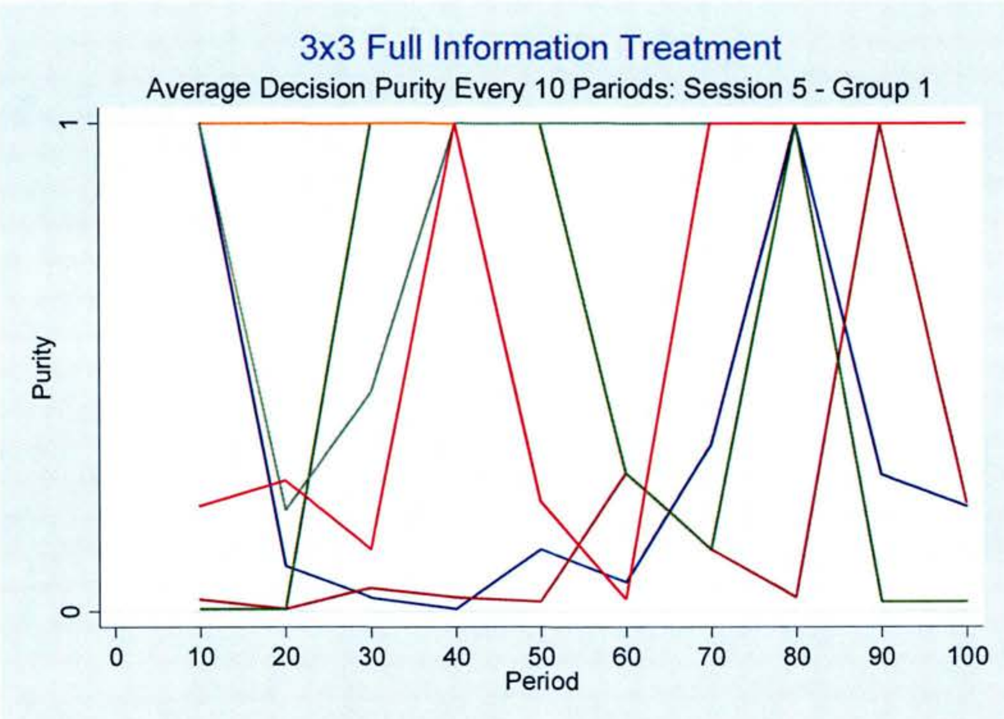

Figure 3.95: 3 × 3 Full Information Treatment (Session 5, Group 1): Average Individual Decision Purity Every 10 Rounds

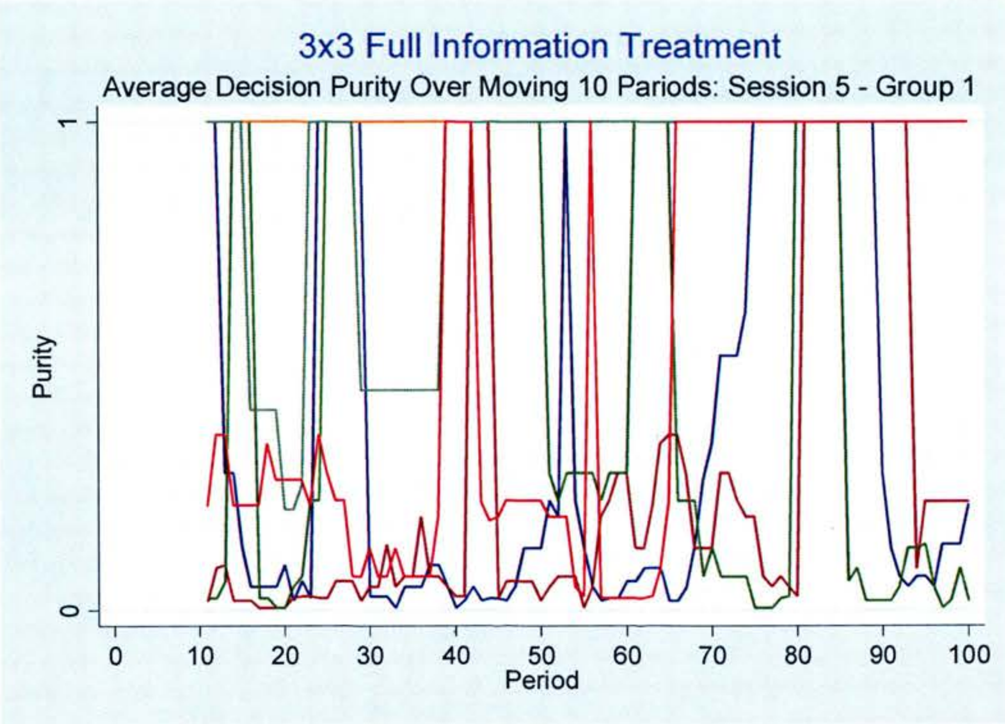

Figure 3.96: 3 × 3 Full Information Treatment (Session 5, Group 1): Average Individual Decision Purity Over Moving 10 Rounds

Session 5, Group 2

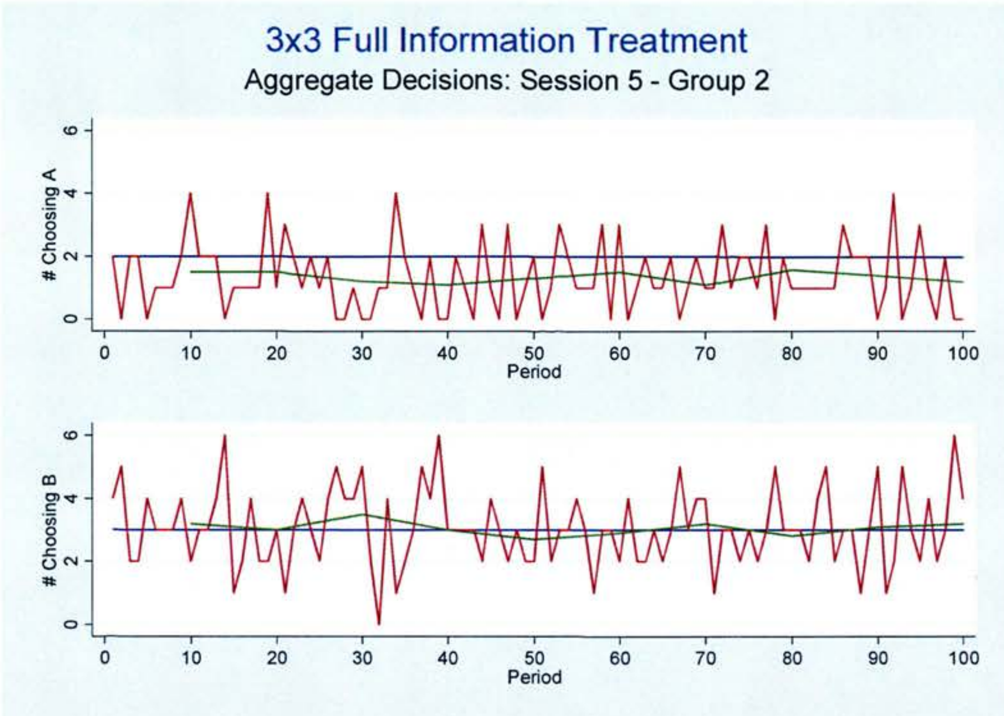

Figure 3.97: 3 × 3 Full Information Treatment (Session 5, Group 2): Aggregate Decisions

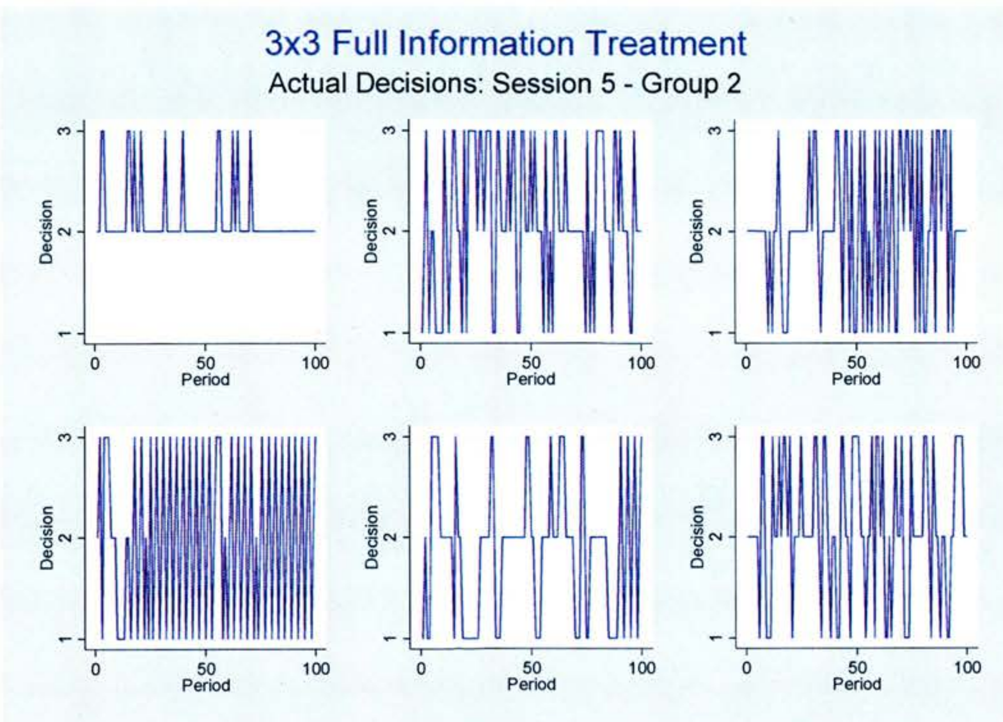

Figure 3.98: 3 × 3 Full Information Treatment (Session 5, Group 2): Individual Decisions

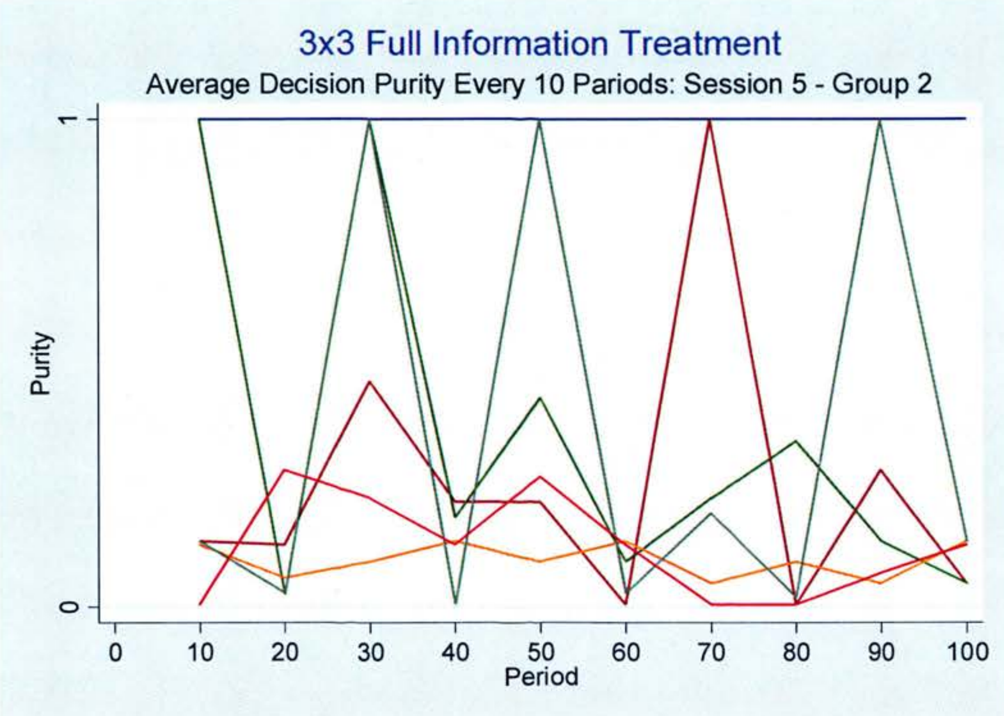

Figure 3.99: 3 × 3 Full Information Treatment (Session 5, Group 2): Average Individual Decision Purity Every 10 Rounds

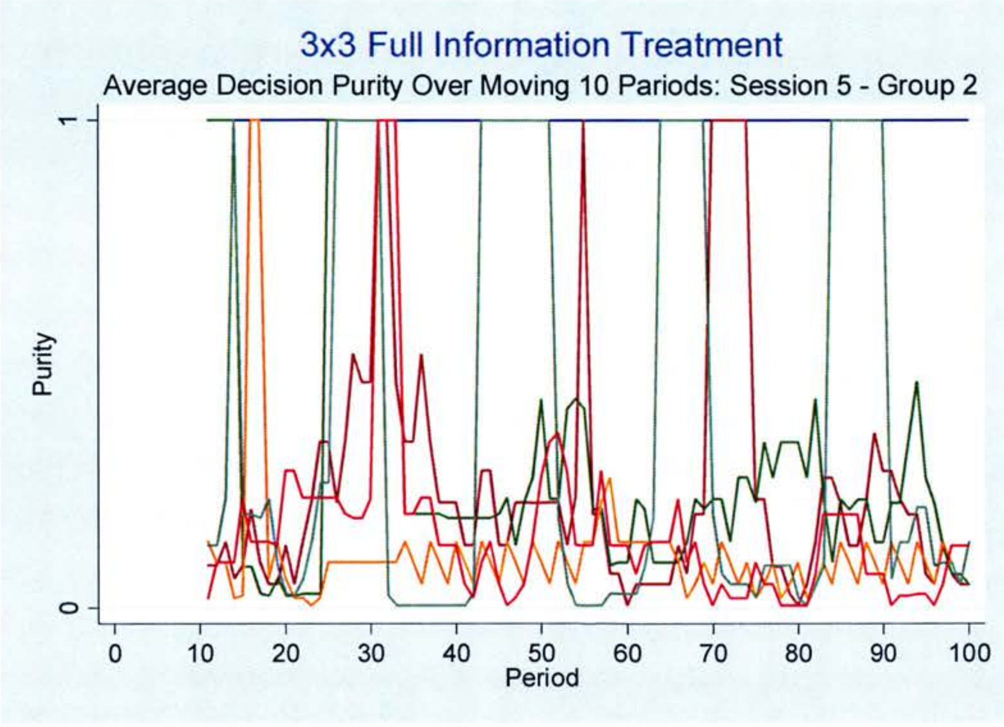

Figure 3.100: 3 × 3 Full Information Treatment (Session 5, Group 2): Average Individual Decision Purity Over Moving 10 Rounds

Session 5, Group 3

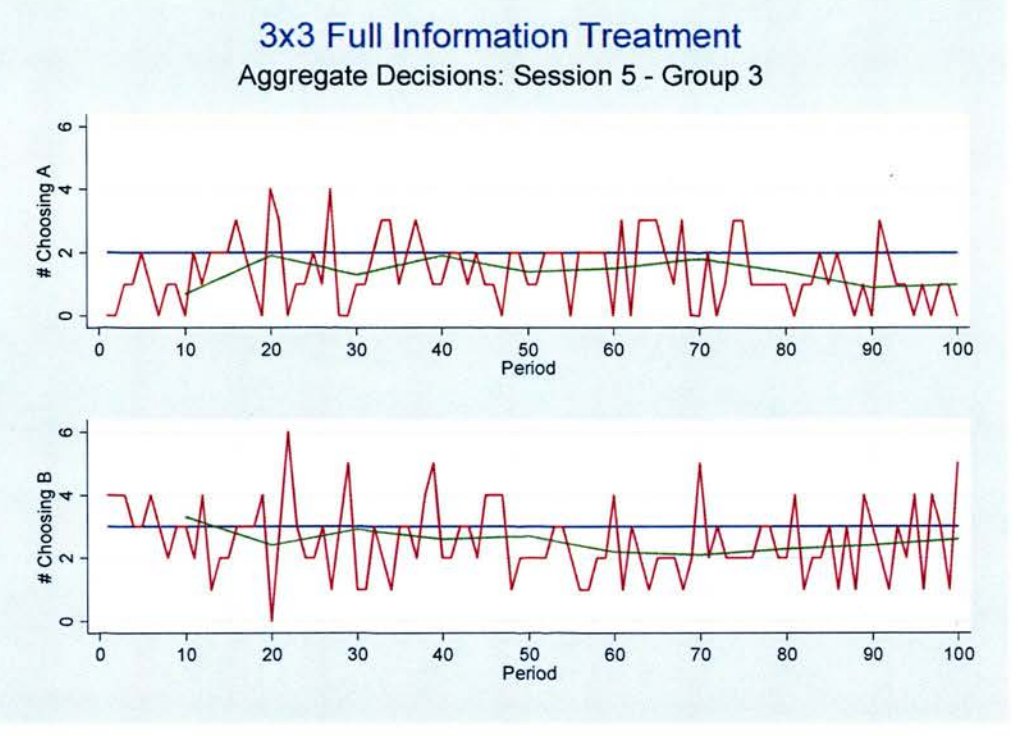

Figure 3.101: 3×3 Full Information Treatment (Session 5, Group 3): Aggregate Decisions

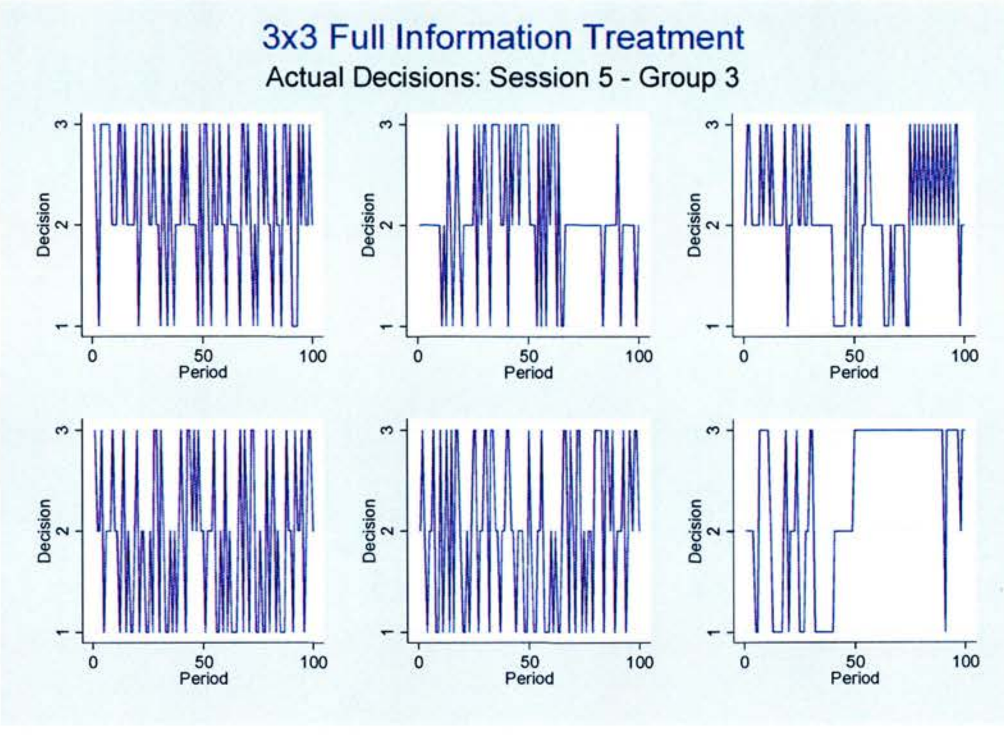

Figure 3.102: 3 × 3 Full Information Treatment (Session 5, Group 3): Individual Decisions

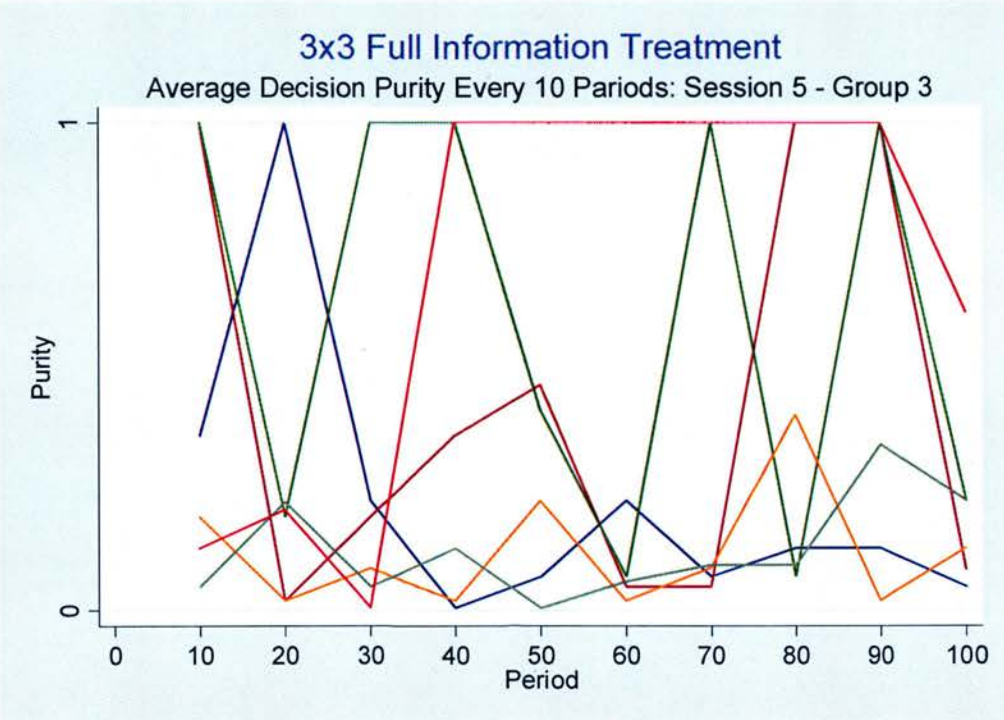

Figure 3.103: 3 × 3 Full Information Treatment (Session 5, Group 3): Average Individual Decision Purity Every 10 Rounds

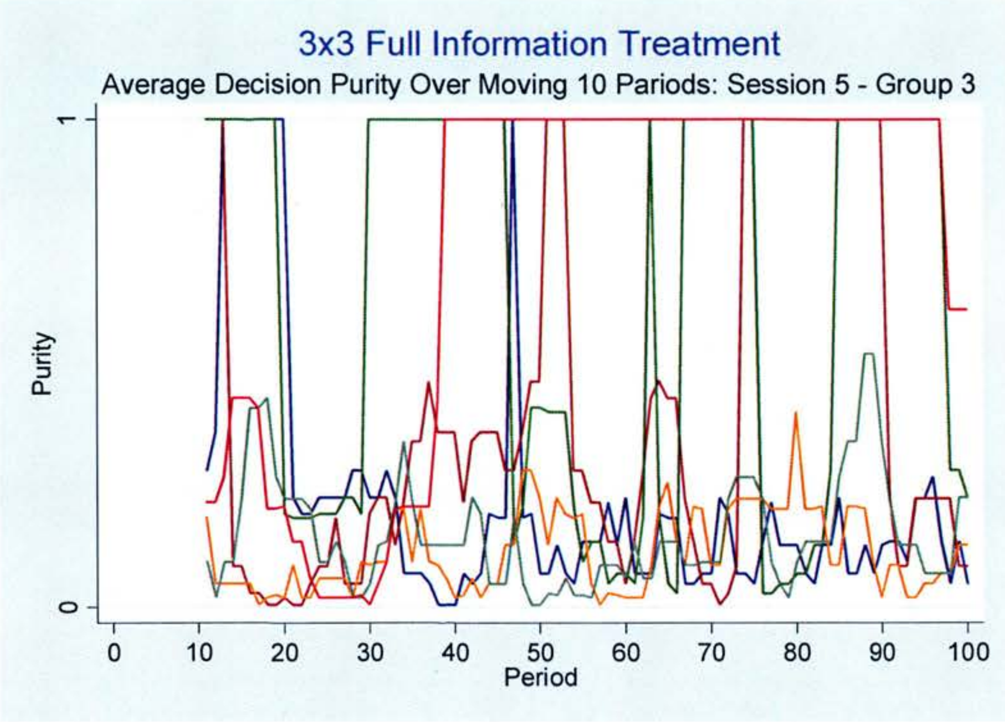

Figure 3.104: 3 × 3 Full Information Treatment (Session 5, Group 3): Average Individual Decision Purity Over Moving 10 Rounds

Appendix 3.B Introductory Script

3.B.1 Session Introduction

Welcome. We thank you for coming here to participate in this session. My name is Duncan Whitehead. I am here to answer your questions. Please feel free to ask any question during the session.

At the end of today's session, you will all receive $10 for showing up. In addition to the $10, you will have the opportunity to earn more money, depending on the outcome of the activity that will be played out in this session.

Now let me tell you a little about our research project. We are interested in how people make decisions. Thus, we will ask you to make a few decisions. All these decisions are individual – that means that you should take your decisions individually.

The total amount of money that you will win at the end of the session is yours to keep. We will pay you this money in addition to the $10 for showing up, but it should be clear that this money is not ours. The money comes from the Social Sciences and Humanities Research Council of Canada so that we can run this study.

Before we start, we would like to make a few points clear:

1. First, this is a study about how people make individual decisions. We ask that you not look at or copy what others are doing. Please do not speak during

this session, and do not tell others what decision you would make. This is very important and we ask you to obey this rule.

2. Second, please listen carefully to all that we say during the session. We will explain to you all that you will be required to do. In addition, I will respond to all your questions during the entire session.

3. Finally, remember that you should always keep with you the number that we gave you when you signed in. Keep it with you at all times. If you lose this number, we will be unable to pay you.

We will now explain to you what you will be doing. Please pay very close attention.

(PROCEED WITH THE TREATMENT INSTRUCTIONS)

3.B.2 Introduction to Consent Forms

Now that we've explained to you what you need to do, you must decide whether you will stay and participate or if you wish to withdraw. If you choose to withdraw, you will receive the $10 show up fee, but you lose the opportunity to win more money in this experiment. If you choose to stay and participate, you will receive the $10 show up fee PLUS whatever you win in the course of the experiment.

Once again, allow me to thank you for coming here today to participate in this session. Please find the two copies of the Consent Form on your desk. We will now review the contents of this document together.

The front side of the form provides general information about the study.

- The title of this study,
- The names and contact details of the study's investigators. Please feel free to contact us at any time regarding any aspect of this study.
- A description of the study. Today's session will last 2 hours. You will not be asked to fill out any questionnaire during today's session.
- The risks and benefits to you of participating in this study. There are no risks or direct benefits to you from participating in this study.

- The costs and payments of participating in this study. There are no costs to participation and you will receive payments in the form of the $10 show up fee plus whatever you win in the course of the experiment.

On the reverse side you will find important statements with respect to:

- Confidentiality. Information will be handled in a confidential manner.
- Right to Withdraw. You may withdraw from this study at any time.
- Voluntary Consent. This is your agreement to participate in the study.

The primary reason for asking you to sign this sheet is that it remains clear that you have chosen to participate in this session, VOLUNTARILY. We wish to protect your interests. That's why we need your signature to verify that you wish to be part of this study. We will keep the form that you've signed, but we will also sign a copy for you to take home with you.

If you agree to the terms in the consent form, please sign where indicated. One copy is yours, the other is ours. I will collect the forms now.

(COLLECT SIGNED FORMS)

3.B.3 Begin Session Script

We are now ready to start the session.

When you have finished making all your decisions, that is, once you have chosen your preferred option in each round and have been informed of your average points earnings over the ten selected periods and your final payment, you will then be directed to collect your earnings. At this point, we will pay you. Once you have been paid, you will be asked to sign a receipt. Once signed, you are free to go.

Feel free to raise your hand if you have a question during the session. We will do our best to answer them. However, please remember that these decisions are personal, so that we cannot help you choose among the options. Remember – it's YOUR decision!

Are there any more questions?

We are now ready. Please turn to your computer screens ready to begin making your decisions.

(BEGIN SESSION)

Appendix 3.C Treatment Instructions

3.C.1 2×2 Aggregate Information Treatment Instructions

Let us begin with a general overview of today's session.

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions you might earn a considerable amount of money that will be paid to you in cash at the end of the session. If you have a question at any time, please feel free to ask the experimenter. Once again we ask that you do not talk with one another for the duration of the session.

Each of you is seated at a desk with a computer workstation, a blank receipt, two blank consent forms, and a pen. You will use the computer workstation to enter individual decisions when prompted to do so. When you have made a decision, you will be given some feedback on the decision you made. Once you have made all your decisions, you will be informed of the payment you will receive, instructed to fill in your receipt, and then directed to collect your earnings for today's session.

Today's session will progress as follows:

1. At the beginning of today's session the computer will allocate you at random and anomalously into groups. You will remain in these groups for the duration of today's session.

2. Then in each round you will be matched at random and anomalously with another individual from your group.

3. Once matched, you will be asked to make a choice over several alternatives. You will earn points that are jointly determined by your decision and the decision of the individual who you are matched with.

4. Once all members of your group have made their decisions you will move onto the next round. Here you will be informed of your decision, the decision of the individual who you were matched with, and the points you received in the previous round. You will also be provided with some information as to the decisions made by your group as a whole.

5. Once you have made all your decisions, you will be informed of the payment you will receive at the end of today's session.

Now let us look today's session in more detail.

Let us begin with the group allocation. There are 18 participants in today's session. At the beginning of the session you will be allocated to one of 3 groups by the computer. Each group will have exactly 6 members. Throughout today's session you will remain in these initially allocated groups. The group allocation and the identity of the individual group members will be anonymous.

There are 100 rounds in today's session. In each round, you will be randomly and anonymously matched with another individual from your group. You will be informed for which round you are making a decision for on the top left hand side of you screen. Once all players have made their decisions the computer will proceeds to the next round.

How *do you input your decisions?*

In each round you will be asked to make a choice between two alternatives, either action **A** or action **B**. You enter your decision by using the mouse to place the curser over either button **A** or **B** in the Decision box on the right hand side of your screen and clicking. Please be aware that once you click on one of these buttons your choice for that round is final. Once all players have entered their decisions the experiment will precede the next round.

How are *your points earned in the round determined?*

The decision you make and that of the individual who you were matched with determine the points you earn in the round. On the board behind me you will see a table depicting your points earned in each round given the choice of action you and your opponent make.

	A	B
A	80	260
B	360	120

- If you choose action **A** and your opponent chooses action **A**, your points earned for the round is **80**.
- If you choose action **A** and your opponent chooses action **B**, your points earned for the round is **260**.
- If you choose action **B** and your opponent chooses action **A**, your points earned for the round is **360**.
- If you choose action **B** and your opponent chooses action **B**, your points earned for the round is **120**.

This information shall stay up on the board for your reference throughout today's session.

How your earnings are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the points your earned resulting from your decision in the previous round. Remember the points you earn in each round are jointly determined by your decision and the decision of the individual you are matched with. You will be informed of the both your decision, the decision of the individual you were matched with and the points you earned in each previous round in the Personal History box on the left hand side of your screen.

How Group decisions are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the decisions made by all members of your group (including yourself) in the Group History box at the bottom of your screen. You will be informed of the total and the share of individuals who choose each action in each previous round.

How is you final take home payment calculated?

Following the completion of the 100 rounds your take home payment will then be calculated. The payment you will receive at the end of the session will be equal to the average number of the points you received in 10 random selected rounds times the exchange rate of 1 Canadian dollar for every 10 points. The computer selects 10 rounds at random for each participant - each round has an equal chance of being selected. You will be informed of your payment in the Payment box at the end of the session.

Once you have been informed of your earnings, you then will be directed to collect your earnings. At this point, we will pay you. Once you have been paid, you will be asked to sign a receipt. Once signed, you are free to go.

3.C.2 3×3 Aggregate Information Treatment Instructions

Let us begin with a general overview of today's session.

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions you might earn a considerable amount of money that will be paid to you in cash at the end of the session. If you have a question at any time, please feel free to ask the experimenter. Once again we ask that you do not talk with one another for the duration of the session.

Each of you is seated at a desk with a computer workstation, a blank receipt, two blank consent forms, and a pen. You will use the computer workstation to enter individual decisions when prompted to do so. When you have made a decision, you will be given some feedback on the decision you made. Once you have made all your decisions, you will be informed of the payment you will receive, instructed to fill in your receipt, and then directed to collect your earnings for today's session.

Today's session will progress as follows:

1. At the beginning of today's session the computer will allocate you at random and anomalously into groups. You will remain in these groups for the duration of today's session.

2. Then in each round you will be matched at random and anomalously with another individual from your group.

3. Once matched, you will be asked to make a choice over several alternatives. You will earn points that are jointly determined by your decision and the decision of the individual who you are matched with.

4. Once all members of your group have made their decisions you will move onto the next round. Here you will be informed of your decision, the decision of the individual who you were matched with, and the points you received in the previous round. You will also be provided with some information as to the decisions made by your group as a whole.

5. Once you have made all your decisions, you will be informed of the payment you will receive at the end of today's session.

Now let us look today's session in more detail.

Let us begin with the group allocation. There are 18 participants in today's session. At the beginning of the session you will be allocated to one of 3 groups by the computer. Each group will have exactly 6 members. Throughout today's session you will remain in these initially allocated groups. The group allocation and the identity of the individual group members will be anonymous.

There are 100 rounds in today's session. In each round, you will be randomly and anonymously matched with another individual from your group. You will be informed for which round you are making a decision for on the top left hand side of you screen. Once all players have made their decisions the computer will proceeds to the next round.

How *do you input your decisions?*

In each round you will be asked to make a choice between two alternatives, either action **A**, action **B** or action **C**. You enter your decision by using the mouse to place the curser over either button **A**, **B** or **C** in the Decision box on the right hand side of your screen and clicking. Please be aware that once you click on one of these buttons your choice for that round is final. Once all players have entered their decisions the experiment will precede the next round.

How are *your points earned in the round determined?*

The decision you make and that of the individual who you were matched with determine the points you earn in the round. On the board behind me you will see a table depicting your points earned in each round given the choice of action you and your opponent make.

	A	B	C
A	80	200	260
B	480	120	180
C	160	280	100

- If you choose action **A** and your opponent chooses action **A**, your points earned for the round is **80**.
- If you choose action **A** and your opponent chooses action **B**, your points earned for the round is **200**.
- If you choose action **A** and your opponent chooses action **C**, your points earned for the round is **260**.
- If you choose action **B** and your opponent chooses action **A**, your points earned for the round is **480**.
- If you choose action **B** and your opponent chooses action **B**, your points earned for the round is **120**.
- If you choose action **B** and your opponent chooses action **C**, your points earned for the round is **180**.

- If you choose action **C** and your opponent chooses action **A**, your points earned for the round is **160**.
- If you choose action **C** and your opponent chooses action **B**, your points earned for the round is **280**.
- If you choose action **C** and your opponent chooses action **C**, your points earned for the round is **100**.

This information shall stay up on the board for your reference throughout today's session.

How your earnings are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the points your earned resulting from your decision in the previous round. Remember the points you earn in each round are jointly determined by your decision and the decision of the individual you are matched with. You will be informed of the both your decision, the decision of the individual you were matched with and the points you earned in each previous round in the Personal History box on the left hand side of your screen.

How Group decisions are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the decisions made by all members of your group (including yourself) in the Group History box at the bottom of your screen. You will be informed of the total and the share of individuals who choose each action in each previous round.

How is you final take home payment calculated?

Following the completion of the 100 rounds your take home payment will then be calculated. The payment you will receive at the end of the session will be equal to the average number of the points you received in 10 random selected rounds times the exchange rate of 1 Canadian dollar for every 10 points. The computer selects 10 rounds at random for each participant - each round has an equal chance of being selected. You will be informed of your payment in the Payment box at the end of the session.

Once you have been informed of your earnings, you then will be directed to collect your earnings. At this point, we will pay you. Once you have been paid, you will be asked to sign a receipt. Once signed, you are free to go.

3.C.3 2×2 Full Information Treatment Instructions

Let us begin with a general overview of today's session.

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions you might earn a considerable amount of money that will be paid to you in cash at the end of the session. If you have a question at any time, please feel free to ask the experimenter. Once again we ask that you do not talk with one another for the duration of the session.

Each of you is seated at a desk with a computer workstation, a blank receipt, two blank consent forms, and a pen. You will use the computer workstation to enter individual decisions when prompted to do so. When you have made a decision, you will be given some feedback on the decision you made. Once you have made all your decisions, you will be informed of the payment you will receive, instructed to fill in your receipt, and then directed to collect your earnings for today's session.

Today's session will progress as follows:

1. At the beginning of today's session the computer will allocate you at random and anomalously into groups. You will remain in these groups for the duration of today's session.

2. Then in each round you will be matched at random and anomalously with another individual from your group.

3. Once matched, you will be asked to make a choice over several alternatives. You will earn points that are jointly determined by your decision and the decision of the individual who you are matched with.

4. Once all members of your group have made their decisions you will move onto the next round. Here you will be informed of your decision, the decision of the individual who you were matched with, and the points you received in the previous round. You will also be provided with some information as to the decisions made by your group as a whole.

5. Once you have made all your decisions, you will be informed of the payment you will receive at the end of today's session.

Now let us look today's session in more detail.

Let us begin with the group allocation. There are 18 participants in today's session. At the beginning of the session you will be allocated to one of 3 groups by the computer. Each group will have exactly 6 members. Throughout today's session you will remain in these initially allocated groups. The group allocation and the identity of the individual group members will be anonymous.

There are 100 rounds in today's session. In each round, you will be randomly and anonymously matched with another individual from your group. You will be informed for which round you are making a decision for on the top left hand side of you screen. Once all players have made their decisions the computer will proceeds to the next round.

How *do you input your decisions?*

In each round you will be asked to make a choice between two alternatives, either action **A** or action **B**. You enter your decision by using the mouse to place the curser over either button **A** or **B** in the Decision box on the right hand side of your screen and clicking. Please be aware that once you click on one of these buttons your choice for that round is final. Once all players have entered their decisions the experiment will precede the next round.

How are *your points earned in the round determined?*

The decision you make and that of the individual who you were matched with determine the points you earn in the round. On the board behind me you will see a table depicting your points earned in each round given the choice of action you and your opponent make.

	A	B
A	80	260
B	360	120

- If you choose action **A** and your opponent chooses action **A**, your points earned for the round is **80**.
- If you choose action **A** and your opponent chooses action **B**, your points earned for the round is **260**.
- If you choose action **B** and your opponent chooses action **A**, your points earned for the round is **360**.
- If you choose action **B** and your opponent chooses action **B**, your points earned for the round is **120**.

This information shall stay on the board for your reference throughout today's session.

How your earnings are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the points your earned resulting from your decision in the previous round. Remember the points you earn in each round are jointly determined by your decision and the decision of the individual you are matched with.

You will be informed of the both your decision, the decision of the individual you were matched with and the points you earned in each previous round in the Personal History box on the left hand side of your screen.

How Group decisions are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the decisions made by all members of your group (including yourself) in the Group History box at the bottom of your screen. You will be informed of the individual decision each member of your group who made in each previous round.

How is you final take home payment calculated?

Following the completion of the 100 rounds your take home payment will then be calculated. The payment you will receive at the end of the session will be equal to the average number of the points you received in 10 random selected rounds times the exchange rate of 1 Canadian dollar for every 10 points. The computer selects 10 rounds at random for each participant - each round has an equal chance of being selected. You will be informed of your payment in the Payment box at the end of the session.

Once you have been informed of your earnings, you then will be directed to collect your earnings. At this point, we will pay you. Once you have been paid, you will be asked to sign a receipt. Once signed, you are free to go.

3.C.4 3×3 Full Information Treatment Instructions

Let us begin with a general overview of today's session.

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions you might earn a considerable amount of money that will be paid to you in cash at the end of the session. If you have a question at any time, please feel free to ask the experimenter. Once again we ask that you do not talk with one another for the duration of the session.

Each of you is seated at a desk with a computer workstation, a blank receipt, two blank consent forms, and a pen. You will use the computer workstation to enter individual decisions when prompted to do so. When you have made a decision, you will be given some feedback on the decision you made. Once you have made all your decisions, you will be informed of the payment you will receive, instructed to fill in your receipt, and then directed to collect your earnings for today's session.

Today's session will progress as follows:

1. At the beginning of today's session the computer will allocate you at random and anomalously into groups. You will remain in these groups for the duration of today's session.

2. Then in each round you will be matched at random and anomalously with another individual from your group.

3. Once matched, you will be asked to make a choice over several alternatives. You will earn points that are jointly determined by your decision and the decision of the individual who you are matched with.

4. Once all members of your group have made their decisions you will move onto the next round. Here you will be informed of your decision, the decision of the individual who you were matched with, and the points you received in the previous round. You will also be provided with some information as to the decisions made by your group as a whole.

5. Once you have made all your decisions, you will be informed of the payment you will receive at the end of today's session.

Now let us look today's session in more detail.

Let us begin with the group allocation. There are 18 participants in today's session. At the beginning of the session you will be allocated to one of 3 groups by the computer. Each group will have exactly 6 members. Throughout today's session you will remain in these initially allocated groups. The group allocation and the identity of the individual group members will be anonymous.

There are 100 rounds in today's session. In each round, you will be randomly and anonymously matched with another individual from your group. You will be informed for which round you are making a decision for on the top left hand side of you screen. Once all players have made their decisions the computer will proceeds to the next round.

How *do you input your decisions?*

In each round you will be asked to make a choice between two alternatives, either action **A**, action **B** or action **C**. You enter your decision by using the mouse to place the curser over either button **A**, **B** or **C** in the Decision box on the right hand side of your screen and clicking. Please be aware that once you click on one of these buttons your choice for that round is final. Once all players have entered their decisions the experiment will precede the next round.

How are *your points earned in the round determined?*

The decision you make and that of the individual who you were matched with determine the points you earn in the round. On the board behind me you will see a table depicting your points earned in each round given the choice of action you and your opponent make.

	A	B	C
A	80	200	260
B	480	120	180
C	160	280	100

- If you choose action **A** and your opponent chooses action **A**, your points earned for the round is **80**.
- If you choose action **A** and your opponent chooses action **B**, your points earned for the round is **200**.
- If you choose action **A** and your opponent chooses action **C**, your points earned for the round is **260**.
- If you choose action **B** and your opponent chooses action **A**, your points earned for the round is **480**.
- If you choose action **B** and your opponent chooses action **B**, your points earned for the round is **120**.
- If you choose action **B** and your opponent chooses action **C**, your points earned for the round is **180**.

- If you choose action **C** and your opponent chooses action **A**, your points earned for the round is **160**.
- If you choose action **C** and your opponent chooses action **B**, your points earned for the round is **280**.
- If you choose action **C** and your opponent chooses action **C**, your points earned for the round is **100**.

This information stay written up on the board for your reference throughout today's session.

How your earnings are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the points your earned resulting from your decision in the previous round. Remember the points you earn in each round are jointly determined by your decision and the decision of the individual you are matched with. You will be informed of the both your decision, the decision of the individual you were matched with and the points you earned in each previous round in the Personal History box on the left hand side of your screen.

How Group decisions are reported

Once all players have made their decisions and the session has proceeded to the next round you will be informed of the decisions made by all members of your group (including yourself) in the Group History box at the bottom of your screen. You will be informed of the individual decision each member of your group who made in each previous round.

How is you final take home payment calculated?

Following the completion of the 100 rounds your take home payment will then be calculated. The payment you will receive at the end of the session will be equal to the average number of the points you received in 10 random selected rounds times the exchange rate of 1 Canadian dollar for every 10 points. The computer selects 10 rounds at random for each participant - each round has an equal chance of being selected. You will be informed of your payment in the Payment box at the end of the session.

Once you have been informed of your earnings, you then will be directed to collect your earnings. At this point, we will pay you. Once you have been paid, you will be asked to sign a receipt. Once signed, you are free to go.

Bibliography

ANDERSON, S. P., J. K. GOEREE, AND C. A. HOLT (2001): "Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium," *Games and Economic Behavior*, 34, 177–199.

ARTHUR, B. W. (1991): "Learning and Adaptive Economic Behavior. Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality.," *American Economic Review*, 81(2), 353–359.

——— (1994): "Complexity in Economic Theory. Inductive Reasoning and Bounded Rationality.," *American Economic Review*, 82(2), 406–411.

——— (1999): "The End of Certainty in Economics," in *Einstein Meets Magritte*, ed. by B. Aerts, and Mathijs. Kluwer Academic Publishers.

ARTHUR, B. W., Y. M. ERMOLIEV, AND Y. M. KANIOVSKI (1983): "A Generalized Urn Problem and its Applications," *Cybernetics*, 19, 61–71.

——— (1985): "Strong Laws for a Class of Path-Dependent Stochastic Processes with Applications," in *Proceedings of the International Conference in Stochastic Optimisation*, ed. by S. Arkin, and Wets, New York. Spronger.

AXELROD, R. (1990): *The Evolution of Cooperation.* Penguin Books, London.

BEGGS, A. W. (2005): "On the Convergence of Reinforcement Learning," *Journal of Economic Theory*, 122(1), 1–36.

BELL, A. M., AND W. A. SETHARES (1999): "The El Farol Problem and the Internet: Congestion and Coordination Failure," Society for Computational Economics: Computing in Economics and Finance Woking Paper, No. 812.

——— (2001): "Avoiding Global Congestion Using Decentralised Adaptive Agents," *IEEE Transactions on Signal Processing*, 49(11), 2873–2879.

BELL, A. M., W. A. SETHARES, AND J. A. BUCKLEW (2003): "Coordination Failure as a Source of Congestion in Information Networks," *IEEE Transactions on Signal Processing*, 51(3), 875–885.

BENAÏM, M. (1999): *Dynamics of Stochastic Approximation Algorithms*, Le Séminaire de Probabilites: Lecture Notes in Mathematics, Vol 1709. Springer-Verlag, New York.

BENAÏM, M., AND J. W. WEIBULL (2003): "Deterministic Approximation of Stochastic Evolution in Games," *Econometrica*, 71(3), 873–903.

BENVENISTE, A., M. MÉTIVIER, AND P. PRIOURET (1990): *Adaptive Algorithms and Stochastic Approximations.* Springer-Verlag, London.

BINMORE, K. (1992): *Fun and Games: A Text on Game Theory.* D. C. Heath and Company, Lexington.

——— (2001): "Evolution and Mixed Strategies," *Games and Economic Behavior*, 34, 200–226.

BLUME, L. E. (1997): "Population Games," in *The Economy as an Evolving Complex System II*, ed. by W. B. Arthur, S. N. Durlauf, and D. A. Lane, chap. 16, pp. 425–460. Addison-Wesley, Reading.

BORKAR, V. S., S. JAIN, AND G. RANGARAJAN (1998): "Dynamics of Individual Specialization and Global Diversification in Communities," *Complexity*, 3(3), 50–56.

——— (1999): "Generalized Replicator Dynamics as a Model of Specialization and Diversity in Societies," *Advances in Complex Systems*, 1(4), 325–359.

——— (2002): "Collective Behaviour and Diversity in Economic Communities: Some Insights from an Evolutionary Game," in *Econophysics: An Emergent Science. Proceedings of the 1st Workshop on Econophysics, Budapest, 1997*, ed. by J. Kertész, and I. Kondor, pp. 267–279, Amsterdam. Kluwer.

BOYLAN, R. T. (1992): "Laws of Large Numbers for Dynamical Systems with Randomly Matched Individuals," *Journal of Economic Theory*, 57(2), 473–504.

CASTI, J. L. (1996): "Seeing the Light at El Farol," *Complexity*, 1(5), 7–10.

CASTI, J. L., AND A. KARLQVIST (1995): *Cooperation and Conflict in General Evolutionary Processes.* Wiley-Interscience Publication, New York.

CHALLET, D. (2004): "Competition Between Adaptive Agents: Learning and Collective Efficiency," in *Collectives and the Design of Complex Systems*, ed. by T. Kagan, and D. Wolpert, chap. 5, pp. 145–160. Springer-Verlag.

CHALLET, D., M. MARSILI, AND G. OTTINO (2004): "Shedding Light on El Farol," *Physica A: Statistical Mechanics and its Applications*, 332(1), 469–482.

CHALLET, D., AND Y.-C. ZHANG (1997): "Emergence of Cooperation and Organization in an Evolutionary Game," *Physica A: Statistical Mechanics and its Applications*, 246(3-4), 407–418.

CHENG, J. Q. (1997): "The Mixed Strategy Equilibria and Adaptive Dynamics in the Bar Problem," Santa Fe Institute: Santa Fe Working Paper, No. 96-08-063.

CHEUNG, Y.-W., AND D. FRIEDMAN (1998): "A Comparison of Learning and Replicator Dynamics Using Experimental Data," *Journal of Economic Behavior and Organization*, 35, 263–280.

CRAWFORD, V. P. (1989): "Learning and Mixed-Strategy Equilibria in Evolutionary Games," *Journal of Theoretical Biology*, 140(2), 537–550.

DUFFY, J., AND E. HOPKINS (2005): "Learning, Information and Sorting in Market Entry Games: Theory and Evidence," *Games and Economic Behavior*, 51, 31–62.

ENGLE-WARNICK, J., AND E. HOPKINS (2006): "A Simple Test of Learning Threory," Centre for Interuniversity Research and Analysis on Organizations: CIRANO Working Papers No. 2006s-30.

EREV, I., AND A. RAPOPORT (1998): "Coordination, "Magic," and Reinforcement Learning in a Market Entry Game," *Games and Economic Behavior*, 23(2), 146–175.

EREV, I., AND A. E. ROTH (1998): "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88(4), 848–881.

FARAGO, J., A. GREENWALD, AND K. HALL (2002): "Fair and Efficient Solutions to the Santa Fe Bar Problem," in *Proceedings of the Grace Hopper Celebration of Women in Computing.*

FISCHBACHER, U. (2007): "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

FRANKE, R. (2003): "Reinforcement Learning in the El Farol Model," *Journal of Economic Behavior and Organization*, 51, 367–388.

FRIEDMAN, D. (1996): "Equilibrium in Evolutionary Games: Some Experimental Results," *The Economic Journal*, 106(434), 1–25.

FUDENBERG, D., AND D. M. KREPS (1993): "Learning Mixed Equilibria," *Games and Economic Behavior*, 5(3), 320–367.

FUDENBERG, D., AND D. K. LEVINE (1998): *The Theory of Learning in Games.* MIT Press.

——— (2008): "Learning and Equilibrium," Harvard University: Working Paper.

FUDENBERG, D., AND S. TAKAHASHI (2008): "Heterogeneous Beliefs and Local Information in Stochastic Fictitious Play," Princeton University: Working Paper.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory.* The MIT Press, Cambridge, Massachusetts.

FUNDENBERG, D., AND D. M. KREPS (1993): "Learning Mixed Equilibria," *Games and Economic Behavior*, 5(3), 320–367.

GABORA, L. (2004): "Ideas are Not Replicators But Minds are," *Biology and Philosophy*, 19(1), 127–143.

GARDNER, R. (1995): *Games for Business and Economics.* John Wiley and Sons, Chichester.

GENIN, D. I., AND E. KATOK (2006): "Non-Equilibrium Dynamics of the Market Entry Game," *submitted to Games and Economic Behavior.*

GIBBONS, R. (1992a): *Game Theory for Applied Economists.* Princeton University Press, Princeton.

——— (1992b): *A Primer in Game Theory.* Pearson Education, Essex.

GOEREE, J. K., AND C. A. HOLT (2002): "An Explanation of Anomalous Behavior in Binary-Choice Games: Entry, Voting, Public Goods, and the Volunteers' Dilemma," University of Virginia, Department of Economics: Virginia Economics Online Papers, No. 328.

GREENWALD, A., B. MISHRA, AND R. PARIKH (1998): "The Santa Fe Problem Revisited: Theoretical and Practical Implications," in *Proceedings of the Festival on Game Theory: Interactive Dynamics and Learning*, Stony Brook. SUNY.

GREINER, B. (2004): "An Online Recruitment System for Economic Experiments," in *Forschung und Wissenschaftliches Rechnen. GWDG Bericht 63.*, ed. by V. M. Kurt Kremer, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.

HARSANYI, J. C. (1973): "Game with Radomly Distributed Payoffs: A New Rationale for Mixed Strategy Equilibrium Points," *International Journal of Game Theory*, 2, 1–23.

HARSANYI, J. C., AND R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games.* The MIT Press, Cambridges, Massasschusets.

HART, S. (2001): "A General Class of Adaptive Strategies," *Journal of Economic Theory*, 98, 26–54.

HART, S., AND A. MAS-COLELL (2000): "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68(5), 1127–1150.

HILL, B. M., D. LANE, AND W. SUDDERTH (1980): "A Strong Law for some Generalized Urn Processes," *Annals of Probability*, 8(2), 214–226.

HOFBAUER, J. (2000): "From Nash and Brown to Maynard Smith: Equilibria, Dynamics and ESS," *Selection*, 1, 81–88.

HOFBAUER, J., AND K. SIGMUND (1998): *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge.

HOLZMAN, R., AND N. LAW-YONE (1997): "Strong Equilibrium in Congestion Games," *Games and Economic Behavior*, 21, 85–101.

HOPKINS, E. (1999): "Learning, Matching, and Aggregation," *Games and Economic Behavior*, 26, 79–110.

——— (2002): "Two Competing Models of How People Learn in Games," *Econometrica*, 70(6), 2141–2166.

HOPKINS, E., AND M. POSCH (2005): "Attainability of Boundary Points under Reinforcement Learning," *Games and Economic Behavior*, 53(1), 110–125.

JOSEF.HOFBAUER, AND E. HOPKINS (2005): "Learning in Perturbed Asymmetric Games," *Games and Economic Behavior*, 52(1), 133–152.

KAGEL, J. H., AND A. E. ROTH (1995): *The Handbook of Experimental Economics*. Princeton University Press, Princeton.

KERTÉSZ, J., AND I. KONDOR (2002): *Econophysics: An Emergent Science. Proceedings of the 1st Workshop on Econophysics, Budapest, 1997.* Kluwer, Amsterdam.

KOJIMA, F., AND S. TAKAHASHI (2007): "Anti-Coordination Games and Dynamic Stability," *International Game Theory Review*, 9(4), 667 – 688.

KUKUSHKIN, N. S. (1999): "Potential Games: A Purely Ordinal Approach," *Economics Letters*, 64(3), 279–283.

LEADY, J. R. (2002): "If Nobody Is Going There Anymore Because It's Too Crowded, Then Who Is Going? Experimental Evidence of Learning and Imitation in the El Farol Coordination Game," Ph.D. thesis, University of Michigan.

MASS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*. Oxford University Press, Oxford.

MAYNARD SMITH, J. (1974): "The Theory of Games and Animal Conflicts," *Journal of Theoretical Biology*, 47, 209–221.

MCKELVEY, R. D., AND T. R. PALFREY (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

MILCATAICH, I. (1996): "Congestion Games with Player-Specific Payoff Functions," *Games and Economic Behavior*, 13, 111–124.

MILGROM, P., AND J. ROBERTS (1990): "Rationalizability, Learning, and Equilibrium in Game with Strategic Complementatities," *Econometrica*, 58(6), 1255–1277.

MONDERER, D., AND L. S. SHAPLEY (1996): "Potential Games," *Games and Economic Behavior*, 14(1), 124–143.

MORGAN, J., AND M. SEFTON (2002): "An Experimental Investigation of Unprofitable Games," *Games and Economic Behavior*, 40, 123–146.

MORRIS, S., AND T. UI (2005): "Generalized Potentials and Robust Sets of Equilibria," *Journal of Economic Theory*, 124(1), 45–78.

NASH, J. (1951): "Non-Cooperative Games," *Annals of Mathematics*, 54(2), 286–295.

OSBORNE, M. J., AND A. RUBINSTEIN (1994): *A Course in Game Theory.* MIT Press, London.

PARKES, D. C., AND D. STEINIG (1997): "The Santa Fe Bar Problem: Achieving Robust Coordination with Randomization and Simple Myopic Learning," University of Pennsylvania: Working Paper.

PEMANTLE, R. (1990): "Nonconvergence to Unstable Points in Urn Models and Stochastic Approximations," *Annals of Probability*, 18(2), 698–712.

RAPOPORT, A., D. A. SEALE, I. EREV, AND J. A. SUNDALI (1998): "Equilibrium Play in Large Group Market Entry Games," *Management Science*, 44(1), 119–141.

ROBINSON, J. (1951): "An Iterative Method of Solving a Game," *Annals of Mathematics*, 54(2), 296–301.

ROSENTHAL, R. W. (1973): "A Class of Games Possessing Pure-Strategy Nash Equilibria," *International Journal of Game Theory*, 2, 65–67.

ROTH, A. E., AND I. EREV (1995): "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economics Behavior*, 8, 164–212.

RUBINSTEIN, A. (1998): *Modeling Bounded Rationality.* MIT Press, London.

SAMUELSON, L. (1997): *Evolutionary Games and Equilibrium Selection.* The MIT Press, Cambridge, Massachusetts.

SANDHOLM, W. H. (1998): "Simple and Clever Decision Rules for a Model of Evolution," *Economic Letters*, 61(2), 165–170.

——— (2001): "Potential Games with Continuous Player Sets," *Journal of Economic Theory*, 97, 81–108.

——— (2005): "Negative Externalities and Evolutionary Implementation," *Review of Economic Studies*, 72(3), 885–915.

SELTEN, R., AND W. GÜTH (1982): "Equilibrium Point Selection in a Class of Market Entry Games," in *Games, Economic Dynamics and Time Series Analysis*, ed. by M. Deistler, E. Fürst, and G. Schwödiauer, chap. Equilibrium point selection in a class of market entry games. Physica-Verlag, Vienna.

UI, T. (2000): "A Shapley Value Representation of Potential Games," *Games and Economic Behavior*, 31(1), 121–135.

——— (2001): "Robust Equilibria of Potential Games," *Econometrica*, 69(5), 1373–1380.

VEGA-REDONDO, F. (1996): *Evolution, Games, and Economic Behaviour.* Oxford University Press, Oxford.

WEIBULL, J. W. (1995): *Evolutionary Game Theory.* MIT Press, London.

——— (1998): "What Have We Learned from Evolutionary Game Theory so Far," The Research Institute of Industrial Economics: IUI Working Paper Series, No. 487.

ZAMBRANO, E. (2004): "The Interplay Between Analytics and Computation in the Study of Congestion Externalities: The Case of the El Farol Problem.," *Journal of Public Economic Theory*, 6(2), 375–395.

ZWICK, R., AND A. RAPOPORT (2002): "Tacit Coordination in a Decentralized Market Entry Game with Fixed Capacity," *Experimental Economics*, 5, 253–272.