An Experimental Investigation of Stochastic Adjustment Dynamics*

Wooyoung Lim† Philip R. Neary‡

September 3, 2016

Abstract

This paper describes an experiment designed to test which, if any, stochastic adjustment dynamic most accurately captures the behaviour of a large population. The setting is a large population coordination game in which two different groups have differing preferences over equilibria. We find that subject behaviour is highly consistent with the myopic best-response learning rule with deviations from this rule that are (i) dependent on the myopic best-response payoff but not on the deviation payoff, and (ii) directed in the sense of being group-dependent. We also find a time trend to deviations, with the magnitude tapering off as time progresses. This is in contrast to much of the theoretical literature that supposes a variety of other specifications of learning rules and both time-independent and payoff-dependent explanations for deviations.

Keywords: Stochastic Adjustment Dynamics; Experiment; The Language Game; Evolutionary Game Theory.

JEL Classification: C72, C73, C92.

*Special thanks to Vince Crawford, Xun Lu, and Jonathan Newton for detailed comments, and to Yong Kil Ahn for excellent research assistance. We would also like to thank Jesper Bagger, Chris Gee, Sotiris Georganas, Jacob Goeree, Sung-Ha Hwang, Philippe Jehiel, Michihiro Kandori, Heinrich Nax, Peter Neary, Santiago Oliveros, Juan Pablo Rud, Hamid Sabourian, Ija Trapeznikova, and seminar participants at the North-American ESA Conference, NUI Maynooth, Royal Holloway, and Universität St. Gallen. An associate editor and two anonymous referees gave much constructive criticism that improved the paper immeasurably. Remaining errors are ours.

†*Email*: wooyoung@ust.hk; *Web*: http://wooyoung78lim.wix.com/wooyounglim. *Address*: Department of Economics, HKUST, Clear Water Bay, Kowloon, Hong Kong.

‡*Email*: philip.neary@rhul.ac.uk; *Web*: https://sites.google.com/site/prneary/. *Address*: Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX.

1 Introduction

While the concept of Nash Equilibrium is by far the most ubiquitous tool for analysing strategic situations, there are still many settings in which making a confident prediction is difficult. Given this, much focus has shifted to non-equilibrium concepts of *learning*, modelled as *stochastic adjustment dynamics*.[1] In this paper we describe a laboratory experiment designed to test which, if any, stochastic adjustment dynamic best captures the bahaviour of a large population in a particular strategic environment.

There are two behavioural components of any stochastic adjustment dynamic that need to be identified. First it is assumed that players follow a simple updating rule, or *learning rule*, whenever afforded a revision opportunity. Second, players are not cast as infallible, but rather assumed to occasionally deviate from the learning rule by choosing an action not prescribed by it. This is often explained as subjects making 'mistakes', 'errors', 'trembling', or even as experiencing temporary (but un-modeled) preference shocks. Pinning down the above is of huge practical importance, as once aggregated from the individual level, even minute differences in each of the two ingredients listed above can propel population behaviour towards very different outcomes.

Identifying the manner in which subjects deviated from their learning rule is the main focus of our study.[2] Of course, a deviation can only be defined relative to a learning rule so identifying the learning rule was our first task. In doing this we found that more than 92% of choices taken were a myopic best-response, hereafter mBR, to the previous period's population behaviour. For agents who follow the mBR learning rule there are three 'standard' models of deviations that appear in the literature: (i) payoff-independent deviations (Kandori, Mailath, and Rob, 1993; Young, 1993) are precisely that - equally likely no matter what the current circumstances; (ii) payoff-

[1] Fudenberg and Levine (1998) and Young (2005) are textbook treatments.

[2] Via administered exit surveys from the experiments, we learned that some deviations were intended as a signal to others of their unhappiness with the current state. Given that deviations seem to have several sources - some of which are not really mistakes at all, but rather forward-looking attempts to improve future payoffs - we use the term "deviation" throughout. We thank an Associate Editor for suggesting this.

dependent deviations (Blume, 1993) are parameterised by the logit function and, other than the parameterisation, are also self explanatory; (iii) directed deviations (Naidu, Hwang, and Bowles, 2010) postulate that deviations are intentional, and as such will only occur in certain ways (formally, intentional deviations are modelled by restricting the support of the deviations at certain states). Across all our treatments, surprisingly we found that subjects deviated in a manner dependent upon the payoff to the mBR but not upon the payoff to the deviation. We also found a time trend to deviations with the magnitude tapering off as time progresses, and also evidence that mistakes are directed.[3]

Our experiment involved 200 rounds of a large population coordination game, the *Language Game* of Neary (2012). The Language Game seems well suited to pin down deviations from the mBR learning rule. It is a coordination game in which actions are strategic complements where two homogeneous groups have differing preferences. It is similar to a large population Battle of the Sexes in that each group has a preferred action and these differ across group, but it is richer as each 'sex' interacts not only with those from the other group but also with those from their own group. That is, there are three kinds of interaction: 'men' with 'men', 'men' with 'women', and 'women' with 'women'.[4] The tension in preferences coupled with the fact that everyone's payoff depends on the same global environment means that at any given profile, each of the three deviation models listed above inclines the different groups to behave in identifiably different ways.[5] Lastly, the Language Game is not a complex environment, so it is both

[3]Each of the standard models of deviations are time-homogeneous though there are some that have a time component (see for example Pak (2008) and Robles (1998)).

[4]In this sense it can also be interpreted as a heterogeneous *contracting game* (Young, 1998), where each player is contracting with both those with similar preferences and those with opposing preferences. Other large population models with heterogeneity appear in Friedman (1998), Myatt and Wallace (2004), and Sandholm (2007). Oprea, Henwood, and Friedman (2011) contrasts asymmetric contests with homogeneous models. They consider two treatments of a symmetric hawk-dove game. One treatment has a one-population matching protocol where long run play conforms to the unique symmetric mixed strategy equilibrium. The other treatment has two-populations, and in this instance behaviour converges to an asymmetric pure strategy equilibrium.

[5]Since subjects interact with both those in their own group and with those in the other group, a non-mBR can even serve as a (costly) signalling device to fellow group members in the event that a 'undesirable' outcome has been reached. Furthermore, again to do with the structure of the game,

extremely accessible for subjects and easily codeable, and as such we are not compelled to use a 'watered down' version of it.

There is a large literature on what equilibria may emerge in the *long run*. In particular, stochastic adjustment dynamics can select equilibria - the so called 'stochastically stable equilibria' (Foster and Young, 1990) - when a game is repeated indefinitely.[6] This predominantly theoretical literature has been used to discuss how norms and conventions may develop (Lewis, 1969; Young, 1996) and to discuss how and why certain standards may emerge while others fail (Arthur, 1989). While the empirical testing of a theoretical concept requiring an infinite horizon is impossible, our results may still have something to contribute. If we can better understand the mechanics by which a system proceeds towards the long run, we may be better able to infer where this long run will be.

Our results may also contribute to the critique of stochastic stability due to Bergin and Lipman (1996). They showed that any equilibrium can be selected for an appropriately defined model of deviation, and as such argued that the "nature of the mistake [deviation] process must be analysed more carefully to derive some economically justifiable restrictions ... It is an open question whether and what kinds of interesting restrictions will emerge". That we find deviations dependent on the mBR payoff but not on the deviation payoff is at odds with much of the literature, in particular with the concept of Quantal Response Equilibria introduced in McKelvey and Palfrey (1995).[7]

The classic paper on experimental tests of coordination games is Van Huyck, Battalio, and Beil (1990). They studied behaviour in a symmetric "minimum effort game" and found that the inferior, risk-dominant equilibrium consistently emerged once the population size exceeded a threshold. One issue with this homogeneous environment

directed deviations would always ensure that a particular equilibrium emerges. These features are explained in detail using Figure 2 in subsection 2.2.

[6]While stochastic stability requires infinite time, Ellison (1993) noted that convergence times can vary a lot. Convergence times on networks are typically much faster than in a uniform random matching environment (Montanari and Saberi, 2010; Kreindler and Young, 2013; Azomahou and Opolot, 2014a,b), which is related to the point made first by Ellison (2000) and Samuelson (1994).

[7]Maruta (2002), Yi (2009) and Yi (2011) are evolutionary models with payoff-dependent deviations.

is that all the standard stochastic adjustment dynamics propel behaviour towards this outcome, so while the results are both interesting and robust, it is less clear what the underlying learning mechanism is. Crawford (1991) illustrates that the Stag Hunt is analogous to a 2-action minimum effort game, implying that the Language Game has the property that each group's max action is the other group's min action. It is for this reason that the different groups are pulling in different directions.

Three papers are important to mention. The first and closest, Mäs and Nax (2016), analyses individual level data from a large population experiment in a setting that is formally equivalent to the Language Game being played on a network. The two main differences are that their subjects were not informed as to the structure of the network nor the payoffs of the other participants[8], and that half of their treatments have a payoff range with only 6 elements, so the detection of payoff dependent deviations seems less robust. In contrast, all of our treatments have 80 realised payoff levels (40 for each group), so mapping the 'cost' of a deviation to its 'likelihood' can be more thoroughly explored due to the greater variation. The second paper, Benndorf, Martínez-Martínez, and Normann (2016), also has two populations interacting in a global soup, but instead of actions being strategic complements, all pairwise interactions occur via hawk-dove games. Another interesting feature is that the experiments are run in continuous time The last such paper, Crawford (1995), introduces adaptive dynamics that fuse both rules governing strategy updating and the priors with which players 'enter' the game. That is, it presents a theory encompassing both initial behaviour and repeated behaviour.

The balance of the paper is organised as follows. In the next section, we formally define the Language Game and provide insight as to how the different models of deviations pull in different directions. Section 3 describes the experimental design, while Section 4 presents and discusses the results. Section 5 concludes.

[8]While formally this means that a complex Bayesian Game is being played, it has the nice feature that they can consider the 'uncoupled' and 'completely uncoupled' dynamics of Hart and Mas-Colell (2003) and Foster and Young (2006).

2 The Experiment

In subsection 2.1 we formally introduce the Language Game of Neary (2012), while in 2.2 we use the state space associated with one of our treatments to illustrate how, conditional on the mBR learning rule, different models of deviations behave quite differently.

2.1 The Language Game

The Language Game is a simultaneous-move game defined as the tuple $\{\mathcal{N}, \Pi, S, \mathbb{G}\}$, where $\mathcal{N} := \{1, \ldots, N\}$ is the *population* of players; $\Pi := \{A, B\}$ is a partition of $\mathcal{N}$ into two nonempty homogeneous *groups* A and B of sizes N^A and N^B respectively ($N^A, N^B \geq 2$); $S := \{a, b\}$ is the binary action set common to all players; $\mathbb{G} := \left\{G^{AA}, G^{AB}, G^{BB}\right\}$ is a collection of *local interactions*, where G^{AA} is the pairwise exchange between a player from Group A and a player from Group A, etc. These local interactions are given in Figure 1 below, where $\alpha, \beta \in (1/2, 1)$. Utilities are computed as the average of payoffs earned from playing the field, where the same action must be used with one and all.

$$G^{AA} \qquad\qquad \begin{array}{cc|c|c|} & & \multicolumn{2}{c}{A_2} \\ & & a & b \\ \hline A_1 & a & \alpha, \alpha & 0, 0 \\ \hline & b & 0, 0 & 1-\alpha, 1-\alpha \\ \hline \end{array}$$

$$G^{BB} \qquad\qquad \begin{array}{cc|c|c|} & & \multicolumn{2}{c}{B_2} \\ & & a & b \\ \hline B_1 & a & 1-\beta, 1-\beta & 0, 0 \\ \hline & b & 0, 0 & \beta, \beta \\ \hline \end{array}$$

$$G^{AB} \qquad\qquad \begin{array}{cc|c|c|} & & \multicolumn{2}{c}{B} \\ & & a & b \\ \hline A & a & \alpha, 1-\beta & 0, 0 \\ \hline & b & 0, 0 & 1-\alpha, \beta \\ \hline \end{array}$$

Figure 1: The three local interactions, G^{AA}, G^{AB}, and G^{BB}

Fixing an order on the players, with each Group A player listed before any from Group B, define $\mathbf{S} := \prod_{i=1}^{N^A} S \times \prod_{j=1}^{N^B} S$, with typical element $\mathbf{s}$. When a player chooses action s, from his perspective action profile $\mathbf{s} \in \mathbf{S}$ can be viewed as $(s; \mathbf{s})$. For any population profile $\mathbf{s} \in \mathbf{S}$, let $n_a(\mathbf{s})$ denote the number of players choosing action a at $\mathbf{s}$. (Clearly then the number of players choosing action b at $\mathbf{s}$ is equal to $N - n_a(\mathbf{s})$.) With this notation, the utility a player in group $K \in \Pi$ receives from taking action $s \in \{a, b\}$ when the population profile is $\mathbf{s}$, written $U^K(s; \mathbf{s})$, is given by[9]

$$U^A(a; \mathbf{s}) := \frac{1}{N-1}\Big(n_a(\mathbf{s}) - 1\Big)\alpha \tag{1}$$

$$U^A(b; \mathbf{s}) := \frac{1}{N-1}\Big(N - n_a(\mathbf{s}) - 1\Big)(1-\alpha) \tag{2}$$

$$U^B(a; \mathbf{s}) := \frac{1}{N-1}\Big(n_a(\mathbf{s}) - 1\Big)(1-\beta) \tag{3}$$

$$U^B(b; \mathbf{s}) := \frac{1}{N-1}\Big(N - n_a(\mathbf{s}) - 1\Big)\beta \tag{4}$$

By Theorem 1 in Neary (2012), the only candidates for pure strategy equilibria are the *group-symmetric* profiles $(\mathbf{a}, \mathbf{a}), (\mathbf{a}, \mathbf{b})$, and $(\mathbf{b}, \mathbf{b})$, where the first boldface symbol refers to the action commonly chosen by those in Group A and the second to that chosen by everyone in Group B. While $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$ are always equilibria, profile $(\mathbf{a}, \mathbf{b})$ is an equilibrium if and only if the smaller group's preferences are sufficiently strong. Profiles $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$ are by definition Pareto efficient. Whenever profile $(\mathbf{a}, \mathbf{b})$ is Pareto efficient it must be an equilibrium though the reverse implication does not hold (Theorem 6 in Neary (2012)). Profile $(\mathbf{b}, \mathbf{a})$ is never an equilibrium nor Pareto efficient.

In words, the Language Game is a binary-action coordination game with two homogeneous groups and a tension in preference across group. While this reads similar to the description of a large population Battle of the Sexes, the Language Game is richer as it also allows for own-population effects. The added richness expands the set of

[9]These utility functions also represent the case of random matching though in random matching each player only has 3 possible realised payoffs. While this is a very different interpretation, if players are assumed risk neutral as is standard, optimising behaviour is the same in both.

pure strategy equilibria. That is, while A players prefer the equilibrium where everyone choose a, and B players prefer the equilibrium where everyone choose b, for a large set of parameter configurations the partially-coordinated outcome where A players choose a and B players choose b is also a pure strategy equilibrium. In all of our treatments, this "partially coordinated" outcome is an equilibrium.

2.2 Behaviour in Periods 2-200

We now consider how players make choices in periods 2 and onwards. (For how subjects might approach the Language Game given that it is a novel game in the literature, i.e., how they might behave in Period 1, see Section A of the Online Appendix.) Before choosing their action for next period, the subjects were presented with summary information pertaining solely to the period just passed; information pertaining to any previous period disappeared.

Any stochastic adjustment dynamic is composed of a deterministic component generated by the learning rule and a noise component due to the deviations. The dynamic results from aggregating the behaviour of all individuals. If players never deviated from their learning rule, population behaviour is given by an *adjustment dynamic*. If players deviate from their rule, this adds noise to the deterministic adjustment dynamic so that population behaviour is given by a *stochastic adjustment dynamic*. Across all our treatments the individual learning rule is extremely consistent with mBR which generates a dynamic referred to in the literature as the "best-reply" dynamic. There are three models of deviations that are considered when players follow mBR. These are,

D1 : payoff-independent deviations

D2 : payoff-dependent deviations

D3 : directed deviations

Using an example, we now sketch how each of the three stochastic adjustment dynamics generated by the mBR learning rule and deviation models D1-D3 apply in the Language Game. The parameters used are those of our first treatment, Game1: $(N^A, N^B) = (11, 9)$ and $(\alpha, \beta) = (0.57, 0.67)$. For these parameters, action profile $(\mathbf{a}, \mathbf{b})$ is an equilibrium. We begin first by considering the deterministic (noise-free) best-reply dynamic.

A diagram is helpful. While the Language Game has a population of size N (set equal to 20 in each of our treatments), there are in actuality only two types of player. Thus, at any point in time the action profile can be summarised by a 2-dimensional vector (n_a^A, n_a^B), where n_a^A is the number of players in Group A currently using action a, and n_a^B is the corresponding statistic for Group B. Clearly the sum of n_a^A and n_a^B is n_a as in Equations (1)-(4). Figure 2 shows a condensed version of the action space $\mathbf{S}$, commonly referred to as the *state space*, for our first treatment Game 1.[10] (The state space for Treatments 2 and 3 are illustrated in Section B of the Online Appendix.) The state space is depicted as a 12×10 lattice, with $n_a^A \in \{0, \ldots, 11\}$ on the horizontal axis, and $n_a^B \in \{0, \ldots, 9\}$ on the vertical axis. Each state is depicted by a circle with equilibrium states depicted by large circles. For these parameters, profile $(\mathbf{a}, \mathbf{b})$, identified uniquely with state $(11, 0)$, is an equilibrium.

States are colour coded blue, black, and red. At each of the blue states, action b is the best-response for everybody in the population. At each of the red states, action a is the best-response for all. The black states are those at which there is a tension: at each of these states action a is the unique best-response for all those in Group A, while action b is the unique best-response for all those in Group B. This colour coding of states partitions the state space into a *preference map*. For enhanced clarity, these regions of preference are separated by coloured lines running at 45 degrees from northwest to southeast. Consider for example the set of states where $n_a = 8$ (the set of blue states immediately to the south west of the blue line) and compare this with the set of states where $n_a = 9$ (the set of black states immediately to the north east of the blue line). Either side of this line, the best-response for Group A players is different. As such,

[10]This approach substantially reduces the complexity of the problem as while the action space $\mathbf{S}$ has 2^N elements, the state space has only $(N^A + 1) \times (N^B + 1)$ elements.

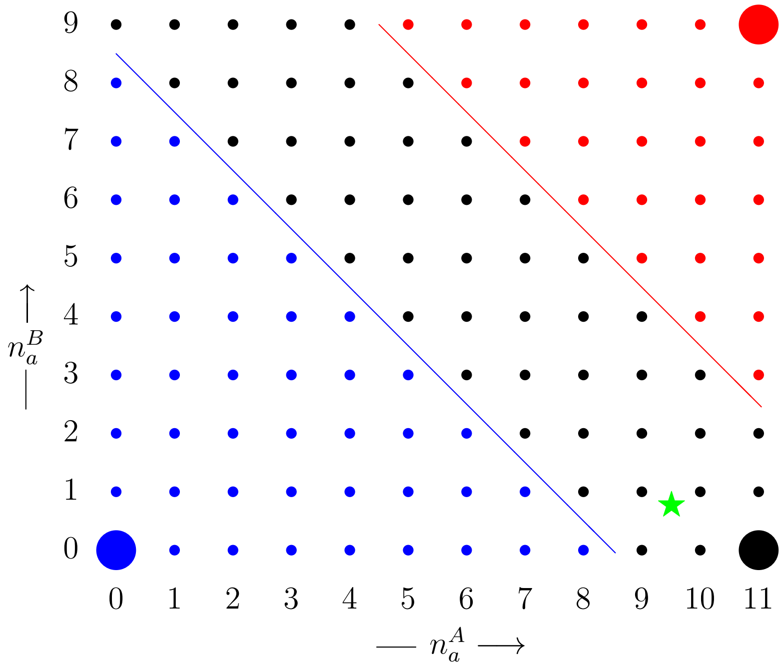


Figure 2: Condensed State Space.

in the language of Schelling (1969, 1971), this line defines *tipping sets* for Group A players.[11]

The relationship between the preference map and the best-reply dynamic is immediate. The colour coding of the preference map conveys which action is a best-response for each group at that state, and the best-reply dynamic implies that everyone takes exactly that prescribed action. As such, whatever the colour of the state in period 1, the following period's state will be the equilibrium of the same colour.[12]

Now let us consider deviations from the best-reply dynamic. As discussed above, the best-reply dynamic is pushing down and left at every blue state, down and right at every black state, and up and right at every red state. Whenever a state that is not an equilibrium state is arrived at, the distance between this state and the intended

[11] The fact that the local interactions are opponent-independent is what makes the boundary states of these sets lie at 45 degrees. If, for example, players had a stronger preference for coordinating with those from their own group, then the boundaries of these sets would be *tilted* away from 45 degrees.

[12] The green star, ★, in Figure 2 denotes the average period 1 population profile for this treatment. As can be seen it is near equilibrium $(\mathbf{a}, \mathbf{b})$, and so quite consistent with our discussion of period 1 behaviour in Section A of the Online Appendix.

equilibrium (as measured according to the L^1 metric) is the number of deviations from mBR that occurred. We now give intuition for how the different deviation models are statistically likely to push the best-reply dynamic in different ways.

D1: Payoff-independent deviations have a clear interpretation. A non-mBR is equally likely at all states for all players independent of Group.

D2: To understand payoff-dependent deviations consider non-equilibrium states $(10, 1)$ and $(10, 3)$ in Figure 2, and consider the perspective of a of Group A player. At both these states action a is his unique best-response. However, the expected payoff from choosing the non-mBR, action b, is higher at state $(10, 1)$ than $(10, 3)$ since he sees more people choosing action b at the latter. So, while a Group A player is unlikely to deviate and choose action b at either state, he is relatively more likely to do so at state $(10, 1)$.

D3: Directed deviations assume that at some states a non-mBR can occur only in certain directions. Formally this is done by restricting the support of the deviation at some states. Given that the Language Game is a binary action game, this means for each group there is a set of states at which they will not deviate. While directed deviations are ad hoc in that the modeller decides which restrictions are/aren't reasonable, we feel there is a natural interpretation in our set up. We propose that directed deviations for Group A players involve choosing action a when action b is the best-response and never choosing action b when action a is the best-response. Similar but opposite for a Group B player. Relating this to the state space in Figure 2, first of all it means that no deviation will ever occur at a state coloured black. Second, it means that only 'rightward' deviations made by Group A players will occur in the blue region and only 'downward' deviations made by Group B players will occur in the red region.

While stochastic stability cannot be tested empirically, the technique does come with a formal language for discussing how easily different equilibria are destabilised that is useful for our analysis. An equilibrium's *basin of attraction* is the set of states from which the dynamics lead to that equilibrium with probability 1. When no deviations occur, so that population behaviour is as described by the best-reply dynamic, then these basins of attraction coincide perfectly with colour coded regions of preference in Figure 2. Now, consider, from each equilibrium the minimum number of deviations needed to reach the basin of attraction of another equilibrium. This number is an equilibrium's *radius* (Ellison, 2000).[13]

The radius provides a useful measure for how robust an equilibrium is to deviations from the mBR learning rule. Transitions against the flow are unlikely and an equilibrium's radius captures this. In the results section, we begin by presenting some coarse population results with transitions between equilibria very apparent.

3 Experimental Design, Hypotheses and Procedure

Now that we have discussed the intuition for the effects of the different models of deviations, we describe our experiment designed to isolate which is most consistent with our subjects' behaviour.

3.1 Design and Hypotheses

All three treatments had a population size of $N = 20$. The treatment variables are the group sizes and the strength of payoffs, so a given game is defined by the 4-tuple $(N^A, N^B, \alpha, \beta)$. Our treatments G1, G2, and G3 are the defined by $(11, 9, 0.57, 0.67)$, $(12, 8, 0.58, 0.71)$, and $(15, 5, 0.58, 0.80)$ respectively.

Before any analysis, we provide some information on individual level payoffs. Table 1 below gives the payoff to each action against every possible population state a Group A subject might face, given that the only information needed is the number of others

[13] With payoff-independent deviations there are lots potential combinations, all requiring 9 people, to escape the symmetric profiles. With payoff dependent deviations, this statement is not entirely accurate as individual deviations may have a different cost depending on the current state. For directed deviations, only transitions along the bottom and right-most boundary can occur.

choosing each action. The first row gives the payoff from choosing action a, the second row the payoff from choosing action b, and the third gives the (absolute value of the) difference between them. Action b is the best-response for all states where $n_a \leq 8$ while action a the best-response for all states with $n_a \geq 9$. The best-response payoff is in red.

$n_a(\mathbf{s})$	0	1	$\cdots$	7	8	9	10	$\cdots$	18	19
$U^A(a; n_a)$	0.000	0.030	$\cdots$	0.210	0.240	0.270	0.300	$\cdots$	0.540	0.570
$U^A(b; n_a)$	0.430	0.407	$\cdots$	0.272	0.249	0.226	0.204	$\cdots$	0.023	0.000
Δ	0.430	0.377	$\cdots$	0.062	0.009	0.044	0.096	$\cdots$	0.517	0.570

Table 1: Payoffs for Group A subject as function of $n_a(\mathbf{s})$ in Game1

Table 1 above provides a perfect template for how a subject using the mBR learning rule should play the game.[14] Table 2 below gives the payoffs at each equilibrium for each of the three treatments. Note in Game 1 and Game 2 that equilibrium $(\mathbf{a}, \mathbf{b})$ is Pareto dominated, whereas in Game 3 it is not.[15]

	Game1		Game2		Game3	
equilibrium	$U^A(\cdot,\cdot)$	$U^B(\cdot,\cdot)$	$U^A(\cdot,\cdot)$	$U^B(\cdot,\cdot)$	$U^A(\cdot,\cdot)$	$U^B(\cdot,\cdot)$
$(\mathbf{a}, \mathbf{a})$	0.57	0.33	0.58	0.29	0.58	0.20
$(\mathbf{a}, \mathbf{b})$	0.30	0.28	0.34	0.26	0.43	0.17
$(\mathbf{b}, \mathbf{b})$	0.43	0.67	0.42	0.71	0.42	0.80

Table 2: Equilibrium Payoffs

[14] Another difference between our design and that of Mäs and Nax (2016) is in the information provided to the subjects. While both games assign payoffs in the manner of Figure 1 and equations (1)-(4), we described the game to subjects in the form of Figure 1 while Mäs and Nax (2016) used that like in Table 1.

[15] While it is true that different stochastic adjustment dynamics propel the system in different directions in the Language Game, this setting also has the interesting feature that for different parameters stochastic stability predictions can also differ across dynamics. When population behaviour is described by the best-reply dynamic and payoff independent deviations, the stochastically stable equilibria are: $(\mathbf{b}, \mathbf{b})$ for Game 1, $(\mathbf{a}, \mathbf{a})$ for Games 2 and 3. When population behaviour is described by the best-reply dynamic and payoff dependent deviations parameterised by the logit function, the stochastically stable equilibria are: $(\mathbf{b}, \mathbf{b})$ for Game 1, $(\mathbf{a}, \mathbf{a})$ for Game 2, and $(\mathbf{b}, \mathbf{b})$ for Game 3. When population behaviour is described by the best-reply dynamic and intentional deviations, equilibrium $(\mathbf{b}, \mathbf{b})$ is stochastically stable for all treatments. It should also be noted that varying the revision protocol can affect the stochastically stable equilibrium when deviations are payoff independent. All these features are discussed in detail in Neary (2013).

Our first hypothesis is very straightforward. In actuality, it says little more than stochastic adjustment dynamics are worthy of consideration in a laboratory setting. Evaluating this hypothesis is simple. All one needs to verify is that, in at least one of our treatments, population behaviour is described by an adjustment dynamic which reaches a profile that is a fixed point, and then moves away from it.

Hypothesis 1. [Stochastic Adjustment Dynamics] Individuals deviate from their learning rule sufficiently regularly that strict equilibria can be escaped from.

Since we are considering only three out of a possible infinite number of learning rules, perhaps individuals are behaving in a manner very different from myopic best-response with noise, and yet population behaviour just so happens to corroborate it in each of our treatments. Analysis of the data at the individual level will allow us to either refute or validate (or at least to not refute) this.

That is, there remains a concern that subjects are not perfectly myopic, and as such, in periods 3 and onwards, are conditioning behaviour using information from more than just the immediately preceding period. This issue is easily checked by comparing actions taken to those prescribed by the mBR learning rule.[16] However, in a given round of play, the feedback provided to subjects pertained only to the immediately preceding period. While it remains possible that our subjects were able to a) recall perfectly information from all previous periods, and b) use this information for strategic purposes, the most recent period is inherently focal.

Hypothesis 2 below asks if deviations from the myopic best-response learning rule depend on payoff differentials. Logit deviations are one (nicely-parameterised) model of deviations allowed under Hypothesis 2, but of course there are others. Our analysis of the individual level data tests for any kind of payoff-dependent deviations.

Hypothesis 2. [Payoff-dependence] The probability of a deviation is higher (lower) when the expected payoff from the non-mBR is higher (lower).

[16]Data analysis shows that in all sessions, the percentage of actions that equate to myopic best-responses exceeds 92.3%, so this appears robust.

Hypothesis 3 below asks if deviations from the myopic best-response learning rule are directed. In our environment, the directedness of the deviations insists that deviations differ across group. When action b is the mBR for everyone, the directed deviations model only allows Group A players to deviate, and vice versa when action a is the mBR for everyone.

Hypothesis 3. [Directedness] When action b is the mBR, the probability of a Group A player deviating is higher than the probability of a Group B player deviating. When action a is the mBR, the probability of a Group B player deviating is higher than the probability of a Group A player deviating.

3.2 Experimental Procedure

There were four sessions run for each of the three treatments. All twelve sessions shared the same procedure. All sessions were conducted in English at the Hong Kong University of Science and Technology (HKUST). A total of 240 subjects (=12 sessions of 20 subjects) were recruited from the undergraduate and graduate populations of the university. No subject had any prior experience with this game.

Subjects entered the lab and each was assigned a private computer terminal.[17] Copies of the experimental instructions were distributed and subjects were given 10 minutes to read them. Communication of any sort between the subjects was forbidden throughout thereby removing coalitional effects as a confounding factor.[18] After reading the instructions, but before commencing the session, the subjects were required to answer a brief questionnaire demonstrating that they understood how payoffs would be assigned each period. No session was to start until all students had responded to each question correctly, although no problems were encountered. Finally, the experimenter read the instructions aloud to ensure that the information included in the instructions, that at this point was verified as understood, was mutual knowledge, and, depending

[17] The computer program was written using z-tree (Fischbacher, 2007).

[18] Newton (2012) is a detailed study of how coalitional behaviour can affect stochastic adjustment dynamics. In summary, it can matter a lot so removing it as a possibility was very important.

on the levels of reasoning employed by the subjects, approaching common knowledge.

Play lasted for 200 periods. At the beginning of each period other than the first, each subject was provided with two pieces of information concerning the previous period's play: (i) the total number of players that had chosen each action, and (ii) his own payoff.[19,20] Each subject was then prompted to select his action for the forthcoming period. The only difference across periods was how much time subjects were given to make a decision. We allowed 30 seconds to choose an action in the initial period, 15 seconds in periods 2-10, and 10 seconds in all periods thereafter (periods 11-200).[21] If a subject did not make a choice within the allowed timeframe, his action from the previous period was carried forward. For more details, see the instructions and the z-tree screenshots in the Online Appendices.

Payoffs were assigned as the average of that earned from playing the field. Payoffs were scaled by a factor of 100. For a given treatment, maximum Group A (B) payoffs were given by 100α (100β), while minimum payoffs were zero in all treatments. The groups were labelled as A and B, but actions a and b were labeled as '#' and '&' respectively, so as to reduce the possibility that *group identity* might increase anchoring on a particular action due to its label.[22]

Take home cash was assigned as the sum of rewards from two randomly chosen rounds plus a HK\$40 show-up fee. Average earnings were HK\$135.7 ($\approx$ US\$17.4), with the range of payoffs given by the interval [HK\$60, HK\$190] ($\approx$ [US\$7.83, US\$24.35]).

---

[19] As can be seen from Equations (1)-(4), each one of these pieces of information is sufficient for a subject to compute the other. Both were provided for the sake of clarity.

[20] Importantly, subjects were not told how the actions taken were distributed across the two groups. This information was withheld to avoid the possibility that this information could be used as an external coordination device. It is also in accordance with a) how the game is defined, and b) how it would be played in a genuinely "large" population.

[21] Together with the limited information feedback provided, the restricted time limit for each round placed a practical restriction on subjects' behaviour, arguably prodding it to be more in line with myopic best-response.

[22] The theory of social identity was initially developed by Tajfel and Turner (1979) in the field of social psychology and is now gaining popularity in economics (Chen and Li (2009) is a recent experimental study). In our setting, group identity means that an individual finds a particular action more attractive than just its payoff consequence. Despite the fact that we labeled actions a and b as # and & respectively, it is not a huge cognitive challenge for a subject to realise that there is a particular action that he would like to see coordination on and as such he may come to identify with that action.

4 Results

We begin with some population level analysis depicted in Figure 3. The Figure contains three sets of four panels labelled (a)-(d), (e)-(h), and (i)-(l), with each panel referencing a different session of that treatment (four panels for each of the three treatments, so twelve panels in total). The information in a given panel is as follows. On the horizontal axis is the period, ranging from 1 to 200, and on the vertical axis is the number of players in a particular subset of the population using action a in a given period. The three subsets are Group A (points connected by the light dashed line); Group B (points connected by the dark dashed line); and the total population (points connected by the solid line). The solid line is simply the "sum" of the other two. Regions of preference in the state space, that correspond to basins of attraction under the best-reply dynamic, are represented via shading, with the basins for the symmetric profiles shaded in light blue and the basin for the equilibrium $(\mathbf{a}, \mathbf{b})$ left unshaded.

So how would behaviour evolve if players followed the mBR learning rule and never erred? There would be path dependence, with the initial state dictating the entire trajectory of population behaviour. Whatever region of the state space the initial state fell in, the following state should be the equilibrium in that region and it should remain that state forever more.

With the above in mind, our first hypothesis, Hypothesis 1, is then trivial to check. As can be seen from simple eyeballing, any of the sessions in any of the treatments confirm that the study of noisy dynamics is of value. Specifically, in each session, more than one strict equilibrium is locked in on and then subsequently escaped from. Related to this, refer back to Figure 2. Note that in all sessions of treatment G1, for which Figure 2 depicts the state space, population behaviour began at an action profile very close to $(\mathbf{a}, \mathbf{b})$. The precise starting profiles were, $(10, 0)$, $(11, 1)$ $(9, 1)$, and $(8, 1)$, which corresponds to an average period 1 population profile in G1 of $(9.5, 0.75)$. This is the state depicted by the green star, ★, in Figure 2. Note that from this state, any noise-free adjustment dynamic derived from mBR learning rule must lock in on population profile $(\mathbf{a}, \mathbf{b})$ and stay there forever.

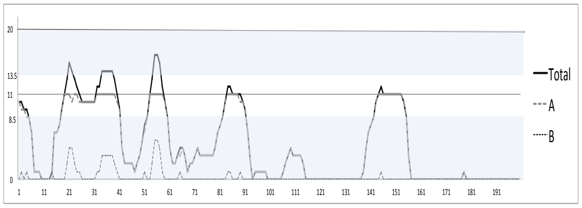
The total number of times that a basin of attraction is escaped from in a given session can be computed from the corresponding panel in Figure 3 by computing the number of times the process 'jumps' across shaded areas. This happened many times in each session. The minimum number of transitions is 5 (G1, Session 2) while the maximum is 27 (G1 Session 4).[23] Such escaping cannot occur with purely deterministic dynamics. Our first result is clear.

Result 1. *Strict equilibria are escaped from, i.e., population behaviour moves between basins of attractions.*

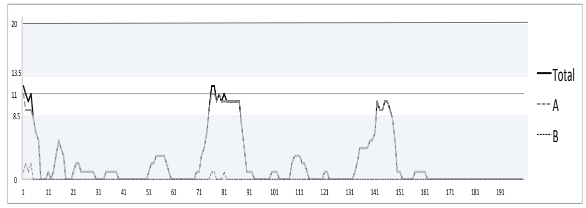
Our main focus in this section is to test if individual deviations can be explained by a logit parameterisation as stated in Hypothesis 2. A deviation should only be classified as such if we first confirm that there is a hard and fast rule that subjects are following. So to analyse how our subjects deviated, we first need to determine precisely what behavioural rule they are following.

As a first pass, consider Figure 4. The figure plots the percentage of actions that were not mBR taken in each round aggregated over all four sessions of all three games. The average deviation percentage aggregated across all rounds were 7.61%, with 10.76% in Game 1, 5.73% in Game 2, and 6.35% in Game 3. The significantly higher likelihood of deviations in Game 1 comes from the outlier (Session 4) of that treatment. The decreasing regularity with which subjects deviate as time progresses is clear, and is evidence that subjects are learning to control their behaviour over time. A Spearman's Rank order test determined the relationship between the average deviation percentage for each round and the round number. The result shows that there is a strong negative monotonic relationship between the two variables (Spearman's $\rho = -0.71$, $p < 0.001$ for all data; $\rho = -0.53$, $p < 0.001$ for Group A; $\rho = -0.75$, $p < 0.001$ for Group B).

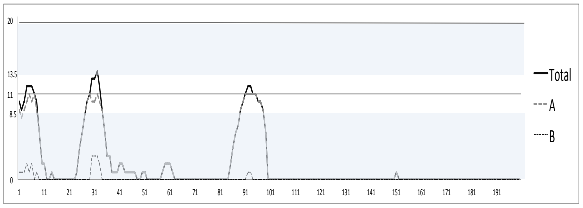
[23] Only in the extremely volatile Session 4 of Treatment G1 was there ever an immediate transition from the basin of attraction of one symmetric equilibrium to the basin of attraction of the other symmetric equilibrium. In all other instances, movements between the two symmetric equilibria involved at least one period where population behaviour was in the basin of attraction of $(\mathbf{a}, \mathbf{b})$.



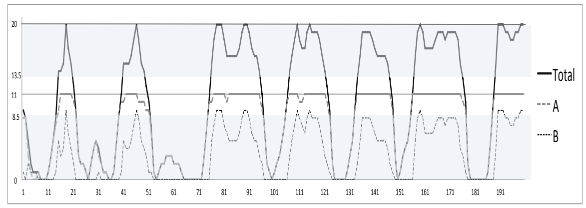
(a) Game 1 Session 1



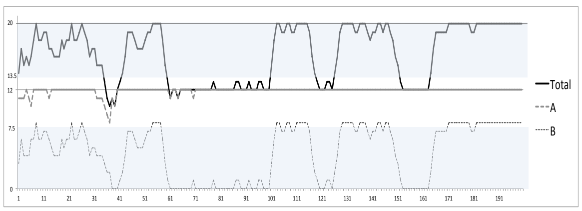
(b) Game 1 Session 2



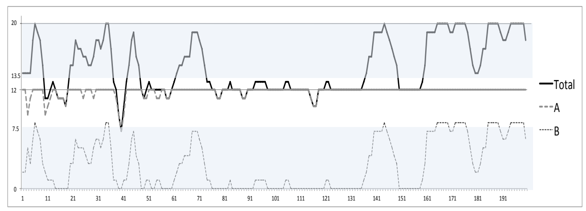
(c) Game 1 Session 3



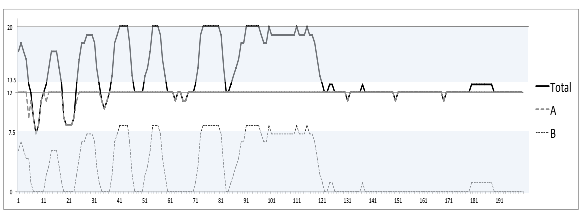
(d) Game 1 Session 4



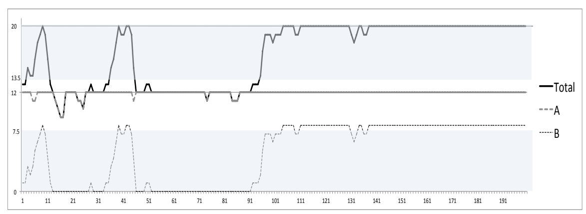
(e) Game 2 Session 1



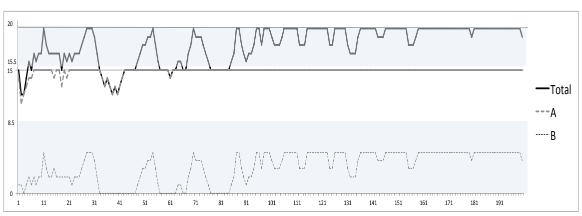
(f) Game 2 Session 2



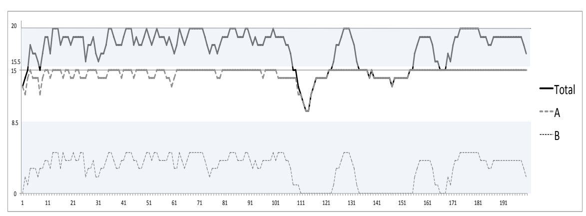
(g) Game 2 Session 3



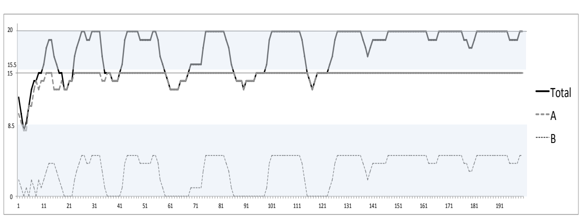
(h) Game 2 Session 4



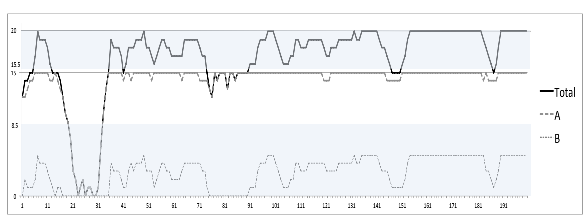
(i) Game 3 Session 1



(j) Game 3 Session 2



(k) Game 3 Session 3



(l) Game 3 Session 4

Figure 3: Trends for frequency of action a

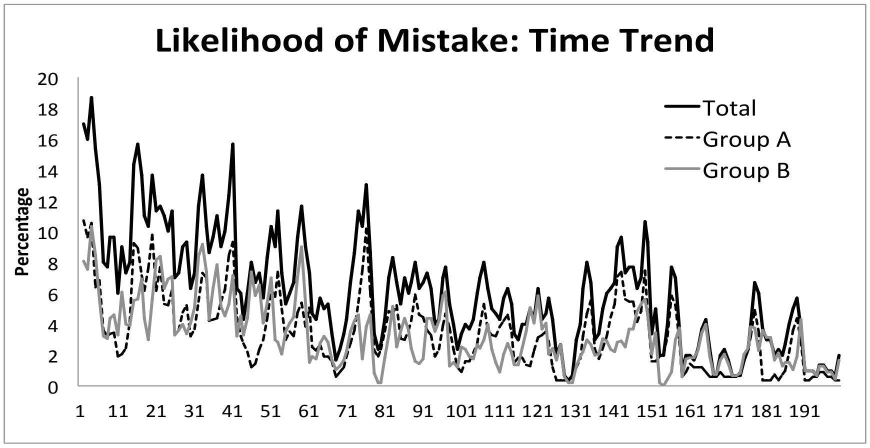
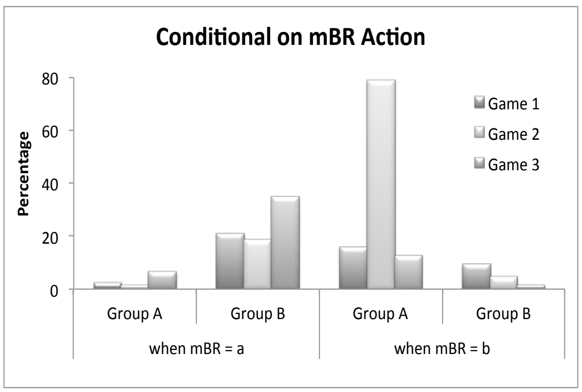
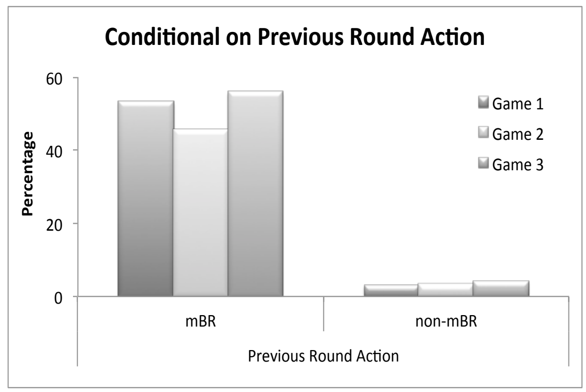


Figure 4: Likelihood of Deviation - Time Trend



(a) Conditional on mBR Action



(b) Conditional on Previous Round Action

Figure 5: Likelihood of Deviations

Figure 5(a) is designed to illustrate whether individual deviations are directed. It plots the average percentages of non-mBR for different groups when mBR is a and when mBR is b. When mBR is a, the likelihoods of a Group A player deviating are 2.25% in G1, 1.38% in G2, and 6.54% in G3, which are significantly and substantially smaller than the likelihoods of a Group B player deviating which were, 20.85% in G1, 18.39% in G2, and 34.75% in G3 (one-sided Wilcoxon signed rank test, $p < 0.001$ for all three pairwise comparisons). Similarly, when mBR is action b, the likelihoods of a Group B player deviating are 9.39% in G1, 4.77% in G2, and 1.47% in G3, which are significantly and substantially smaller than the likelihoods of a Group A player deviating which were, 15.76% in G1, 78.57% in G2, and 12.63% in G3 (one-sided Wilcoxon signed rank test,

$p < 0.001$ for all three pairwise comparisons).

Figure 5(b) illustrates whether individual deviations are intentional, i.e., whether individuals who deviated were *serial deviators* in the next few periods. It plots the percentages of non-mBR taken in each round conditional on non-mBR being taken in the previous round and on mBR being taken in the previous round. Overall, the likelihoods of deviations in each round conditional on the same individual deviating in the previous round are 53.23% in G1, 45.60% in G2, and 55.87% in G3. These numbers are significantly and substantially larger than the likelihoods of a deviation conditional the previous period's behaviour being a mBR which were, 3.05% in G1, 3.23% in G2, and 4.17% in G3 (one-sided Wilcoxon signed rank test, $p < 0.001$ for all three pairwise comparisons).

Based on the findings from the above analyses, which show individual deviations to be time-dependent, directed (group-dependent), and intentional (serially correlated), we conduct the following individual level probit regression for periods $t = 3, \dots, 200$. The dependent variable is M_{it}, and the five regressors were $M_{i,t-1}$, U_{it}^{BR}, U_{it}^{NBR}, t and G_i, where, for time period t and individual i, M_{it} takes the value 1 if the action choice was a non-mBR to the previous period's, population behaviour, and 0 otherwise; G_i is a dummy variable that takes the value 0 if individual i is in Group A, and 1 otherwise; U_{it}^{BR} gives the expected payoff earned from mBR and U_{it}^{NBR} gives the expected payoff earned from the non-mBR.[24] We write ε_{it} for the idiosyncratic error. The coefficients of interest - those on the five regressors above - are given by β_1, β_2, β_3, β_4 and β_5 respectively.

The sign of β_1 reveals the serial correlation between deviations in periods t and $t-1$. The signs of β_2 and β_3 have a straightforward interpretation: if the deviations from the mBR learning rule are dependent on the payoff consequences, then it should be the case

[24] While we used G_i as a dummy variable specifying group identity, we could have defined instead as a variable that takes the value 1 if either [Group is A and mBR $\neq a$] or [Group is B and mBR $\neq B$], and take the value 0 otherwise as such capturing the directedness of a deviation. However, given that in most treatments population behaviour converged to the basin of attraction of a symmetric profile, this specification would be equivalent to capturing group identity.

that the sign of β_2 is negative whereas the sign of β_3 is positive. Coefficients β_4 and β_5 are also uncomplicated: if subjects make fewer deviations over time, then β_4 should be negative. Similarly, if subjects in Group A (Group B) deviate more often than subjects in Group B (Group A), then β_5 should be negative (positive).[25]

Game	Session	# of Obs.	(1) Pre. Deviation Coef.	p-V	(2) U(mBR) Coef.	p-V	(3) U(nmBR) Coef.	p-V	(4) Period Coef.	p-V	(5) Group Coef.	p-V
1	1	3,940	2.256	.000	−.032	.496	−.065	.052	−.006	.000	.523	.470
	2	3,940	2.182	.000	.096	.061	.049	.231	−.005	.000	2.304	.046
	3	3,940	2.229	.000	−.055	.089	−.019	.431	−.002	.024	−.027	.969
	4	3,940	1.451	.000	.022	.562	.496	.136	−.271	.000	1.375	.099
	Total	15,760	2.345	.000	.026	.213	.021	.179	−.004	.000	1.267	.006
2	1	3,940	2.252	.000	.001	.984	.021	.615	−.006	.000	1.445	.037
	2	3,940	2.476	.000	−.000	.999	.046	.216	−.006	.000	1.335	.085
	3	3,940	2.375	.000	−.133	.018	−.095	.059	−.006	.000	−.998	.340
	4	3,940	1.775	.000	.066	.175	.098	.033	−.015	.000	1.693	.062
	Total	15,760	2.434	.000	−.007	.831	.021	.429	−.006	.000	.921	.136
3	1	3,940	1.552	.000	−.101	.000	−.061	.001	−.005	.000	−.266	.029
	2	3,940	1.723	.000	−.110	.210	−.043	.493	−.002	.037	−.801	.000
	3	3,940	1.193	.000	−.131	.000	−.047	.063	−.007	.000	−.648	.000
	4	3,940	.974	.000	−.053	.000	−.007	.207	−.000	.533	−.197	.003
	Total	15,760	1.411	.000	−.062	.000	−.016	.000	−.002	.000	−.137	.002
All Games		47,280	1.826	.000	−.055	.000	−.005	.087	−.003	.000	−.295	.000

Note: Reported p-values are based on robust standard errors clustered on subjects. Abbreviations: # of Obs. = Number of Observations; Pre. Deviation = Previous Period Deviation; U(mBR) = Payoff from mBR action; U(nmBR) = Payoff from the non-mBR; Coef. = Coefficient; p-V = p-value

Table 3: Probit Regression

Table 3 presents the results.[26] Column (1) shows that β_1 is positive and significant at the 1% level for all sessions of all games. The significant serial correlation between deviations in successive periods implies that the departure from the myopic best-response is "intentional". Columns (2) and (3) show that the signs of β_2 and β_3 vary across games and sessions but are more likely to be insignificant. Overall, there is no clear pattern for the two coefficients in terms of the sign and the significance level. Most importantly regarding parameters β_2 and β_3, there is no single session of any game where both $\beta_2 < 0$ and $\beta_3 > 0$ are simultaneously satisfied, as is required for deviations to be consistent with the logit model. Subjects deviate less often as time proceeds.

Compatible with the results from the non-parametric tests, column (5) shows that

[25]The magnitudes of the coefficients do not capture marginal effects because the regression is conducted based on a non-linear model.

[26]The results from running a Logit regression are qualitatively the same and thus not reported here.

deviations are group-dependent in a systematic way. That is, Group B subjects tend to deviate more often than subjects in Group A in G2 and G3 whereas subjects in Group A tend to deviate more often than subjects in Group B in G1. Together with the fact that the profiles to which the population eventually converged is $(\mathbf{a}, \mathbf{a})$ in G2 and G3 and $(\mathbf{b}, \mathbf{b})$ in G1, the systematic group-dependent deviation model implies that the departure from the myopic best-response is "directed". Overall, the result from Games 1-3 are not consistent with the payoff-dependent deviation model with mBR learning rule. Our final result that addresses the issues posed in Hypotheses 2 and 3 can then be summarised as follows:

Result 2. *At the individual level, the probability of a deviation*

1. *is decreasing in the payoff from the best-response (column 2 in Table 3),*
2. *is independent of the payoff from the non-best response (column 3 in Table 3),*
3. *has a time component whereby they decrease in frequency over time (column 4 in Table 3),*
4. *is group-dependent (column 5 in Table 3), and*
5. *is serially correlated (column 1 in Table 3).*

Results 2.1 and 2.2 are clear. We interpret Result 2.3 as subjects learning to control their behaviour over time. Given that deviations don't pay and that the everyone else in the population is making fewer deviations as time progresses, this seems reasonable. We interpret Result 2.4 as deviations being directed, and Result 2.5 as deviations being intentional.

While the evidence for group-dependent, directed deviations is strong, there are many occasions when this is violated. Furthermore, theoretically at least, if deviations occur sufficiently often, then the directed deviation model is unwavering in its insistence that equilibrium $(\mathbf{a}, \mathbf{b})$ be observed a large fraction of the time and this is never the case.

So why are deviations serially correlated? Perhaps, as suggested in Footnote 22 and again related to their intentionality, it may be that some deviations from mBR are signals/attempts to encourage members of one's own population to follow suit. (See again the work of Oprea, Henwood, and Friedman (2011) who find similar behaviour.)

5 Conclusion

This paper describes an experiment whose goal is to determine what stochastic adjustment dynamic best predicts long run behaviour in a large population coordination problem. We use the Language Game of Neary (2012) in which, provided subjects are myopic best-responders, different models of deviations pull the system in directions.

We found that 92% of decisions taken were a myopic best-response to the previous period's population profile. Deviations from this learning rule occurred with a likelihood dependent on the payoff of the myopic best-response, but not with the deviation payoff. We also found autocorrelation in choosing the non-mBR action.

We finish by highlighting some obvious limitations to our set-up. While the issue of which learning model fits best is both an obvious and simple question to ask, we see no reason why there would be a single learning model that would universally dominate all others. So, while our findings seem robust, we caution about over-inferring from them. In our opinion, the better, but more difficult question to ask is how the various aspects of the environment (e.g., the game, the type and framing of feedback provided, the number of periods, etc.) shape the choice of the most 'appropriate' learning model.

References

Arthur, W. B. (1989): "Competing Technologies, Increasing Returns, and Lock-In by Historical Events," *Economic Journal*, 99(394), 116–131.

Azomahou, T. T., and D. Opolot (2014a): "Convergence rates for stochastic evolution in networks," MERIT Working Papers 037, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

——— (2014b): "Epsilon-stability and the speed of learning in network games," MERIT Working Papers 036, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

Benndorf, V., I. Martínez-Martínez, and H.-T. Normann (2016): "Equilibrium Selection with Coupled Populations in Hawk-Dove Games: Theory and Experiment in Continuous Time," *Journal of Economic Theory*, pp. –.

Bergin, J., and B. L. Lipman (1996): "Evolution with State-Dependent Mutations," *Econometrica*, 64(4), 943–956.

Blume, L. E. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387 – 424.

Chen, Y., and S. X. Li (2009): "Group Identity and Social Preferences," *American Economic Review*, 99(1), 431–57.

Crawford, V. P. (1991): "An "evolutionary" interpretation of Van Huyck, Battalio, and Beil's experimental results on coordination," *Games and Economic Behavior*, 3(1), 25–59.

——— (1995): "Adaptive Dynamics in Coordination Games," *Econometrica*, 63(1), 103–143.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica*, 61(5), 1047–1071.

——— (2000): "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67(1), 17–45.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178–178.

Foster, D., and P. Young (1990): "Stochastic evolutionary game dynamics," *Theoretical Population Biology*, 38, 219–232.

Foster, D. P., and H. P. Young (2006): "Regret testing: learning to play Nash equilibrium without knowing you have an opponent," *Theoretical Economics*, 1(3), 341–367.

Friedman, D. (1998): "On economic applications of evolutionary game theory," *Journal of Evolutionary Economics*, 8(1), 15–43.

Fudenberg, D., and D. K. Levine (1998): *The Theory of Learning in Games (Economic Learning and Social Evolution)*. The MIT Press, Cambridge MA.

Hart, S., and A. Mas-Colell (2003): "Uncoupled Dynamics Do Not Lead to Nash Equilibrium," *American Economic Review*, 93(5), 1830–1836.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61(1), 29–56.

Kreindler, G. E., and H. P. Young (2013): "Fast convergence in evolutionary equilibrium selection," *Games and Economic Behavior*, 80(0), 39 – 67.

Lewis, D. K. (1969): *Convention: A Philosophical Study*. Cambridge, Mass: Harvard University Press.

MARUTA, T. (2002): "Binary Games with State Dependent Stochastic Choice," *Journal of Economic Theory*, 103(2), 351 – 376.

MÄS, M., AND H. H. NAX (2016): "A behavioral study of "noise"in coordination games," *Journal of Economic Theory*, 162, 195–208.

MCKELVEY, R. D., AND T. R. PALFREY (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

MONTANARI, A., AND A. SABERI (2010): "The spread of innovations in social networks," *Proceedings of the National Academy of Sciences*, 107(47), 20196–20201.

MYATT, D. P., AND C. C. WALLACE (2004): "Adaptive play by idiosyncratic agents," *Games and Economic Behavior*, 48(1), 124–138.

NAIDU, S., S.-H. HWANG, AND S. BOWLES (2010): "Evolutionary bargaining with intentional idiosyncratic play," *Economics Letters*, 109(1), 31 – 33.

NEARY, P. R. (2012): "Competing conventions," *Games and Economic Behavior*, 76(1), 301 – 328.

NEARY, P. R. (2013): "Supplementing Stochastic Stability," Discussion paper, University of London, Royal Holloway.

NEWTON, J. (2012): "Coalitional stochastic stability," *Games and Economic Behavior*, 75(2), 842–854.

OPREA, R., K. HENWOOD, AND D. FRIEDMAN (2011): "Separating the Hawks from the Doves: Evidence from continuous time laboratory games," *Journal of Economic Theory*, 146(6), 2206 – 2225.

PAK, M. (2008): "Stochastic stability and time-dependent mutations," *Games and Economic Behavior*, 64(2), 650 – 665, Special Issue in Honor of Michael B. Maschler.

Robles, J. (1998): "Evolution with Changing Mutation Rates," *Journal of Economic Theory*, 79(2), 207 – 223.

Samuelson, L. (1994): "Stochastic Stability in Games with Alternative Best Replies," *Journal of Economic Theory*, 64(1), 35–65.

Sandholm, W. H. (2007): "Pigouvian pricing and stochastic evolutionary implementation," *Journal of Economic Theory*, 132(1), 367–382.

Schelling, T. C. (1969): "Models of Segregation," *The American Economic Review*, 59(2), pp. 488–493.

——— (1971): "Dynamic models of segregation," *Journal of Mathematical Sociology*, 1, 143–186.

Tajfel, H., and J. Turner (1979): *An integrative theory of intergroup conflict*.pp. 33–47. Brooks/Cole.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80(1), 234–48.

Yi, K.-O. (2009): "Payoff-dependent mistakes and q-resistant equilibrium," *Economics Letters*, 102(2), 99 – 101.

——— (2011): "Equilibrium Selection with Payoff-Dependent Mistakes," Discussion Paper 1115, Research Institute for Market Economy, Sogang University.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.

——— (1996): "The Economics of Convention," *The Journal of Economic Perspectives*, 10(2), 105–122.

——— (1998): "Conventional Contracts," *The Review of Economic Studies*, 65(4), 773–792.

——— (2005): *Strategic Learning and Its Limits (Arne Ryde Memorial Lectures Series)*. Oxford University Press, USA.