Do repeated game players detect patterns in opponents? Revisiting the Nyarko Schotter belief elicitation experiment

Spiliopoulos, Leonidas

University of Sydney

9 January 2008

Do repeated game players detect patterns in opponents? Revisiting the Nyarko & Schotter belief elicitation experiment

Leonidas Spiliopoulos

University of Sydney

8th January, 2008

Abstract The purpose of this paper is to reexamine the seminal belief elicitation experiment by Nyarko and Schotter (2002) under the prism of pattern recognition. Instead of modeling elicited beliefs by a standard weighted fictitious play model this paper proposes a generalized variant of fictitious play that is able to detect two period patterns in opponents' behavior. Evidence is presented that these generalized pattern detection models provide a better fit than standard weighted fictitious play. Individual heterogeneity was discovered as ten players were classified as employing a two period pattern detection fictitious play model, compared to eleven players who followed a non-pattern detecting fictitious play model. The average estimates of the memory parameter for these classes were 0.678 and 0.456 respectively, with five individual cases where the memory parameter was equal to zero. This is in sharp contrast to the estimates obtained from standard weighted fictitious play models which are centred on one, a bias introduced by the absence of a constant in these models. Non-pattern detecting fictitious play models with memory parameters of zero are equivalent to the win-stay/lose-shift heuristic, and therefore some subjects seem to be employing a simple heuristic alternative to more complex learning models. Simulations of these various belief formation models show that that this simple heuristic is quite effective against other more complex fictitious play models.

JEL classification number: C9; C70; C72; C73

Keywords: Behavioral game theory; Learning; Fictitious play; Pattern detection; Simulations; Beliefs; Repeated games; Mixed Strategy Nash equilibria; Economics and psychology

---

1 Introduction

This paper's main purpose is to test the hypothesis of pattern detection behavior on behalf of subjects engaged in a repeated unique mixed strategy Nash equilibrium game. This will be accomplished by

Correspondence to: lspi4871@usyd.edu.au

reexamining the data from the innovative experiment by Nyarko and Schotter (2002) where players' beliefs about their opponents' actions were directly elicited. Instead of modeling belief formation with a standard weighted fictitious play (henceforth abbreviated to *wfp*) model, a variant of generalized variant of *wfp* will be proposed that will allow for the possibility that players are detecting two-period patterns in their opponents' behavior.

Repeated mixed strategy games have been one of the foci of both the experimental game theory literature and its theoretical counterpart - Camerer (2003) and Kagel (1995) are excellent introductions to the field of behavioral and experimental game theory. The literature is rife with experimental studies investigating issues regarding whether human play is well described by theoretical solutions such the mixed strategy Nash equilibrium (MSNE), the Quantal Response Equilibrium (McKelvey, 1995), or other equilibrium concepts and refinements[1]. These implicitly assume instantaneous equilibration, and therefore remain silent on the learning dynamics of the off-equilibrium path. In response to this, researchers resorted to postulating theories of learning originally inspired by the psychology and artificial intelligence literature which already had a strong history of grappling with such issues. It is generally accepted that in repeated games humans do not initially play according to the MSNE as they are not privy to it, however this raises the issue of whether the MSNE could be learned through experience during the rounds of a repeated game.

It is still debated in the literature as to whether convergence occurs as there are studies coming to conflicting conclusions, and quite often older studies are discredited due to assumptions and the power of statistical tests of convergence to MSNE, see O'Neill (1987) and Brown and Rosenthal (1990). Although the literature on learning in games is primarily concerned with the off-equilibrium path of play of repeated games it can also give new insights as to the attainment (or lack thereof) of a MSNE as the limiting result of the learning process. The behavior of players during the learning phase and their limiting behavior after experience are obviously related and should not be treated separately as the learning rules used by players will influence what the limiting play will be in the long run[2].

Another important role of the experimental branch of this research is to provide empirical justification for the various learning rules posited by theorists, one of which is belief learning. Belief learning assumes that players track the frequency of past actions of their opponents and use this information to form beliefs of their opponent's future play. These beliefs are then translated into final actions through the use of a decision rule. Let a_i represent the action taken by player i. For any action a_j, define player i's count of a_j at time t as:

$$C_i(a_j, t) = \frac{I_{t-1}(a_j) + \sum_{u=1}^{t-2} \gamma_i^u \cdot I_{t-u-1}(a_j)}{1 + \sum_{u=1}^{t-2} \gamma_i^u} \tag{1}$$

where $I_t(a_j)$ is an indicator function which takes the value of one if player j chose action a_j in time period t or the value of zero otherwise.

[1] Examples of specific papers are Palacios-Huerta (2003), Chiappori et al. (2002), Walker and Wooders (2001), Binmore et al. (2001), Bloomfield (1994), Brown and Rosenthal (1990) and Rapoport and Budescu (1997).

[2] Although not examined in this paper, the postulation of learning rules can refine the number of equilibria that are attainable in games with multiple NE, providing an indirect theory of equilibrium selection, as discussed in Haruvy and Stahl (2004).

In standard weighted fictitious play (henceforth abbreviated to *wfp*) (Cheung and Friedman, 1997) the beliefs, $fp1_i(a_j, t)$, of player i regarding the probability of his opponent playing action a_j are equal to the count, $C_i(a_j, t)$[3]. The memory decay parameter for each player is γ_i and memory loss (or weighting of past observations) is assumed to be exponential in discrete time.

Decision rules can be deterministic or stochastic, although evidence from experiments shows that the stochastic approach is more realistic for modeling human behavior as subjects will not necessarily make the same decision each time they are confronted with the same problem. Hence, it is usually assumed that players are probabilistically best responding to the expected payoffs of actions given their beliefs. The decision rule, or probability of playing action i, $P(a_i, t)$, is often a logit function as shown in equation 2, where S_i is the discrete strategy set of player i, and $E(\pi(a_i))$ is equal to the expected payoffs of playing any action a_i given beliefs $fp1_i(a_j, t)$ for all $a_j \in S_j$. The degree of responsiveness to expected payoffs is controlled by the parameter λ_i of the decision rule. As $\lambda_i \to 0$ the probability distribution over actions tends to the uniform distribution where all actions are played with equal probability. However, as $\lambda_i \to \infty$ the decision rule approaches simple deterministic best response, where the action with the highest expected payoff will be played with certainty.

$$P(a_i, t) = \frac{e^{\lambda_i \cdot E(\pi(a_i))}}{\sum_{a_i \in S_i} e^{\lambda_i \cdot E(\pi(a_i))}} \tag{2}$$

The layout of this paper adheres to the following plan. Section 2 initially reviews the paper by Nyarko and Schotter (2002), outlining the experimental procedures followed and their main results. Subsections provide references to the psychology and behavioral economics literature in support of two important premises that motivate the investigation of pattern detection in experimental games. Firstly, economic research will be presented that documents the failure of non-expert human subjects to create random sequences, even if they are consciously trying to do so. Secondly, research from the sequence learning literature in psychology will be presented to advocate that humans are capable of pattern recognition in non-strategic situations. On the basis of these two premises, Section 3 presents the changes and extensions that this paper proposes to the analyses performed by N&S. Attention is paid to the parametric form of the proposed models to capture the possibility of pattern detection and in overcoming some of the problems with the standard *wfp* models used in N&S. Section 4 discusses the econometric techniques adopted in the estimation of the proposed models of behavior. The results of the separate estimation of two proposed models, only one of which allows for pattern detection, are presented in Section 5. Particular attention is paid to the estimated coefficients of the memory parameter $\hat{\gamma}$, which are found to be significantly smaller than the N&S estimates. Section 5.6 allows for the possibility of subject heterogeneity in terms of pattern detection abilities, and classifies each human subject according to the proposed model which best describes their behavior. This will provide an estimate of the frequency of different types of players in the population of experimental subjects, as grouped by their pattern detection abilities. The convergence, or lack thereof, of players' actions to the MSNE prescription is also discussed. Simulations of the different learning models postulated in this paper will be performed in order to investigate the effectiveness and robustness of the said models in terms of payoff accumulation. Finally, Section 6 summarizes

[3] The use of the notation *fp1* to denote standard weighted fictitious play beliefs will become obvious when a two-period pattern detecting variant *fp2* is proposed later on.

the main conclusions of this paper and the appendices discuss some technical details regarding the quadratic scoring rule and the optimization techniques employed in estimating the models.

2 Literature review

Most experimental research in this field collects data on the actual actions chosen by players and then attempts to fit a model of off-equilibrium behavior to this data. This entails the simultaneous estimation of both the belief generating mechanism, which is not directly observable, as well as the decision rule. The concurrent estimation of beliefs and the decision rule however has been shown to be inefficient, and even biased, at estimating the true underlying value parameters. Indeed, estimated parameters sometimes are at opposite extremes of reasonable bounds on parameter values, so that conclusions based on these estimations may not only be inaccurate in terms of magnitude but may even lead to qualitatively erroneous conclusions. Blume et al. (1999) find that identification of true parameter values is poor in experiments with a small numbers of subjects and/or a small number of game rounds. Cabrales and Garcia-Fontes (2000) find similar results as small samples lead not only to inefficient estimates but also to biased estimates of parameters. Salmon (2001) finds problems with the estimation of the EWA model[4], a learning model proposed by Camerer and Ho (1999), as the correct parameters are only recovered correctly roughly half of the time.

2.1 Review of Nyarko and Schotter (2002)

Nyarko and Schotter (2002) made a seminal contribution to the literature by implementing an experimental setup that made beliefs observable, thereby attenuating the econometric identification problems discussed above. In their paper, they not only collect data on the actions of players, but also data about their beliefs by directly asking players to quote them. Before each of their moves, players were asked to state the probability with which they thought their opponents would play their pure strategies and were then requested to state their move for the round. A quadratic scoring rule (QSR) was used as an incentive mechanism for truthful revelation of beliefs. Details of the implications of this incentive mechanism, especially if subjects are not risk-neutral, are provided in Appendix A.

This innovation in the experimental technique allowed N&S to investigate research questions where econometric estimation was problematic due to the unobservability of the underlying beliefs. Firstly, they had data with which to directly examine what the underlying belief formation process really is, as they could now directly try to fit the elicited beliefs. Secondly, they could determine whether or not best responding (albeit with some error) to these elicited beliefs provided a better fit to the data than other learning rules such as reinforcement learning (Roth and Erev, 1995), belief learning (Cheung and Friedman, 1997), or EWA learning (Camerer and Ho, 1999).

2.1.1 Details of the experimental setup The game used in N&S is shown in Table 1, with each subject repeatedly playing the same game 60 times. The mixed strategy Nash equilibrium for both players was to play red 60% of the time and green 40% of the time. They conducted four treatments

[4] Detailed knowledge of the Experience Weighted Attraction (EWA) model is not necessary for this paper, however readers can consult Camerer and Ho (1999) for a detailed discussion.

Table 1 N&S Payoff matrix

		Column player	
		Green	Red
Row player	Green	6,2	3,5
	Red	3,5	5,3

which varied two characteristics of the experimental design: whether beliefs were elicited or not and whether players were randomly rematched after each round of play or not. These treatments are important as they provide baseline comparisons. For example, comparison of the treatments with and without belief elicitation provided evidence that the belief elicitation procedure did not significantly change the subjects' behavior. This is important to test for because if behavior was altered by the act of eliciting beliefs this would reduce the ecological and/or external validity of the experimental results. As this paper is concerned with learning, only the experimental data from the treatment where beliefs were elicited and players had the same opponent over all 60 periods of play are analyzed. Subjects would receive monetary payoffs both according to the accuracy of their beliefs, using an incentive compatible scoring rule, and their payoffs from playing the game.

2.1.2 Summary of results

Weighted fictitious play beliefs are not a good proxy for stated beliefs N&S found that individually estimated *wfp* models, as posited in Equation 1, with a single parameter were actually quite poor at fitting elicited beliefs. Perhaps the most important difference between fitted *wfp* beliefs and the elicited beliefs was the immense difference in variability that the two exhibited. Fictitious play beliefs are very smooth, especially for memory parameters approaching one and large histories of play, whereas elicited beliefs fluctuated wildly from period to period. Indeed, in many cases players stated that they were sure what their opponent would play i.e. stated beliefs equal to zero or one. Also, the variability of beliefs showed no signs of decreasing over the 60 rounds, and contrary to intuition is still significant in magnitude even in the treatment with common knowledge that players are randomly rematched after each round.

An alternative test of how well *wfp* beliefs fit the elicited beliefs is to examine how often the best responses to these beliefs prescribe the same action. In this particular game the mixed strategy equilibrium is 0.6 (for the red action), hence for beliefs of red play between 0.6 and 1, the same action is dictated as the best response, likewise for all beliefs between 0 and 0.6 the other action is the best response. As long as *wfp* beliefs and elicited beliefs are either both greater than 0.6 or both less than 0.6, the prescribed actions will be the same and the *wfp* model will have some credibility in explaining final choices made by players[5]. However, even under this looser requirement *wfp* beliefs and elicited beliefs are on the same side of 0.6 only 65% of the time.

Subjects tend to best respond to stated beliefs rather than weighted fictitious play beliefs Subjects' actions were found to be best responses to elicited stated beliefs 75.7% of the time whereas this

[5] In games with more actions the belief space will be partitioned into more sectors by the MSNE probabilities and therefore model predictions will need to be more accurate in order to lie in the same best response probability interval.

statistic falls significantly to 56.8% when *wfp* beliefs are used instead. This is further indication of the inadequacy of *wfp* beliefs to accurately model stated beliefs.

Fictitious play beliefs are more accurate than stated beliefs in predicting opponents' play N&S reason that even though fictitious play beliefs are more accurate they are not particularly salient for subjects and therefore this may be the reason why they do not switch from an inferior belief formation system to fictitious play.

Learning models that use stated beliefs as inputs to a logistic decision rule are better at fitting data than models using wfp beliefs The finding that stated beliefs are more accurate at predicting actions when the decision rule is a best response process is carried over to a probabilistic decision rule, such as the logistic rule in Equation 2.

Random matching of players after each round does not eliminate the volatility in stated beliefs This counter-intuitive result is surprising as it implies that the volatility of stated beliefs does not arise as a consequence of belief formation given opponents' play since in the random matching case subjects should not condition beliefs on the immediately prior actions of opponents. If anything, they should be using a population fictitious play belief formation process which would entail learning the probabilities of actions across the whole population of players. These beliefs should obviously be very smooth and exhibit very little volatility as there is no reason to condition on previous actions as subjects are not facing the same opponent.

Estimated memory parameters, $\hat{\gamma}$, are very close to values of one The estimated memory parameters are all very close to the value of one, implying there is no memory loss and that all past observations are weighted equally. This is a surprising result as in many other types of experiments humans are found to exhibit less than perfect recall. Section 3.2 discusses how the parametric form of the standard *wfp* model may bias the estimates towards one and proposes a solution to obtain unbiased estimates of $\hat{\gamma}$.

Estimating the memory parameter from action data is problematic Although the estimates of $\hat{\gamma}$ from the belief series were centered on one, the estimates from action data where centered on zero with nine out of 28 estimates returning negative values which clearly have a problematic interpretation. N&S view this as strong evidence that parameter estimates obtained from action data, or in general whenever observable proxies are used as substitutes for unobservable data, should be treated with skepticism.

2.2 Literature review of humans' (in)ability to randomize

The psychology literature has found that when people are asked to create random sequences of variables they are particularly poor at this. Humans tend to over-alternate between actions or regress towards the mean as they find such sequences more representative of distributions. This result is well documented by Bar-Hillel and Wagenaar (1991), Rapoport and Budescu (1997) and Rabin (2002) provides a behavioral model to explain this behavior. For example, when experimental subjects are told to produce random sequences of heads and tails from an unbiased coin, they tend to switch or alternate between heads and tails too often i.e. combinations of heads-heads and tails-tails are each chosen significantly less than 25% of the time. A further consequence of this is that average length of runs (i.e. the number of times in a row heads or tails come up) are too short, or equivalently, that the number of runs in the time series is too high. Also, humans very often try to balance the overall number of heads and tails in the sequences they produce since they see this as being more representative of the distribution they are trying to emulate.

Game theorists have been interested in these documented inefficiencies of human randomization as they imply that one should expect deviations from MSNE i.i.d. behavior or patterns to exist even if humans are capable of solving for the MSNE and try to implement it in game play. In response to many experimental studies that document that human subjects have difficulty in implementing the MSNE, some researchers turned instead to natural field experiments of behavior in games with unique MSNE. Two studies stand out in this literature, that of Palacios-Huerta (2003) which looks at penalty kicks in soccer and that of Walker and Wooders (2001) which examines tennis serves. The former concludes that the minimax hypothesis is confirmed in the case of expert soccer players and that their behavior is serially independent, thereby not exhibiting over- or under-alternation. The latter paper concludes that although there is still evidence of professional players conditioning on past actions, play is closer to the MSNE for professional players or experts than for inexperienced subjects. In conclusion, it appears to be the case that humans are more efficient at randomizing when they have a long history of experience with a specific game and large monetary incentives, as is the case with professional sports players. However, ordinary human subjects who have not repeatedly played exactly the same game for a fair deal of time and do not have large enough monetary incentives to fine-tune their strategies will probably not conform closely to MSNE prescriptions.

2.3 Literature review of pattern detection or sequence learning in humans

Given the evidence presented above, although subjects may be expected to deviate from i.i.d. behavior this will only influence learning if their opponents are able to detect these deviations. Pattern recognition algorithms suffer from increased complexity and memory load, and therefore even if deviations exist they may not be easily detected by humans due to cognitive bounds of memory and information processing. The question of whether humans have the ability to detect patterns in time series sequences is a well established research topic in the psychology literature where it is usually referred to as sequence learning, Clegg et al. (1998) provides a concise introduction. The seminal paper by Nissen and Bullemer (1987) advanced the view that not only does sequence learning exist but it may primarily be an implicit form of learning[6], leading to a flourishing of research in this field. In the current state of the literature, it is widely accepted and beyond dispute that humans exhibit sequence learning and therefore new research is primarily directed at using different experimental methodologies to determine whether sequence learning is explicit or implicit (or perhaps whether both mechanisms coexist).

Before reviewing the evidence from the field of psychology it is necessary to define sequence learning and a measure of the depth of such learning. The definition of nth order probability information is the use of information at time $t-n+1$ to infer behavior at time t. If information

[6] Implicit learning is defined as learning that is not the result of conscious and intentional processes - for an excellent introduction to implicit sequence learning the reader is referred to Cleeremans et al. (1998).

from all periods between $t-n+1$ and $t-1$ are used then this is referred to as nth order adjacent dependency, alternatively if not all of the periods are relevant then it is referred to as non-adjacent dependency[7]. Sequence learning essentially involves pattern detection or recognition because adjacent nth order probability information essentially involves recognizing n consecutive time period strategies. For example, second-order probability information involves calculating the probability of an action conditional on the action played in the previous period. If this probability is very high then the subject has essentially learned that these two temporally consecutive actions are more often than not played in this order i.e. has learned to recognize that this pattern is more common than others.

The existence of sequence learning is well documented by studies such as Remillard and Clark (2001) and Remillard (2007) who find evidence of implicit sequence learning of second- through to fifth-order adjacent and non-adjacent probabilities. Other studies in the experimental psychology field, such as Gomez (1997), have found evidence of explicit knowledge of second-order probabilities in which the subjects were consciously aware of their learning. Another strand of research uses advanced brain scanning techniques to identify the brain structures involved in explicit and implicit sequence learning. Destrebecqz et al. (2003) uncovered a significant correlation between explicit learning exhibited by subjects and activity in the anterior cingulate/mesial prefrontal area of the brain. Aizenstein et al. (2004) discovered that although different activation patterns existed for implicit and explicit learning there was also a large overlap between the regions of the brain that were activated and they conclude that implicit and explicit learning probably occur in parallel in the human brain.

In conclusion, the aforementioned research justifies researching pattern recognition models of learning in game theory as they have documented that such learning is possible in the human brain, whether it be explicit, implicit or a combination of both.

3 Proposals for extending Nyarko and Schotter's analysis

The inadequacy of *wfp* learning rules to detect patterns in opponents' behavior is important as the experimental evidence presented above informs us that humans are capable of some level of implicit and explicit pattern recognition in non-strategic decision making. Therefore it is imperative to investigate whether subjects also perform pattern recognition in strategic games. It should be noted however that in the sequence learning literature the number of rounds is much higher than the rounds in this experiment, or rounds of other game theory experiments. Hence, it is probably not reasonable to expect subjects to be as astute at recognizing patterns as in this literature. We hypothesize that the smaller number of rounds will make higher-order detection such as third- and fourth-order much more difficult, especially because there will not be enough observations on which to base such calculations. Also, subjects will probably be less sensitive to patterns so that the smallest deviation that will be detectable by subjects will be higher than in the psychology literature studies. The magnitude of deviations in game theory experiments may likely be smaller than those in the sequence learning literature, automatically making them tougher to detect. This is because

[7] The numbering of order probabilities in this paper differs from that in the psychology literature for ease of exposition later in the paper. In the psychology literature an nth order probability refers to the use of information at the $t-n$th period instead of the definition given in the main text which refers to the $t-n+1$th period. All references to order probabilities will henceforth refer to the definition adopted in this paper.

the magnitude of deviations in the sequence learning literature is controlled by the experimenter, whereas in human vs. human game theory experiments the magnitude is determined solely by the behavior of subjects. Although sequence learning in the psychology literature occurs over many periods, a significantly large proportion of learning occurs in the early rounds and therefore it is still reasonable to assume that sequence learning will occur during the 100 rounds of this experiment.

The *wfp* learning algorithm used in the literature simply keeps track of the number of times an opponent's action was played in the past and will therefore be able to detect deviations from the MSNE predictions of observed frequencies of single actions, also referred to as first-order probabilities of play. First-order probabilities of play refer to the probability of playing red or green independently of previous play, second-order probabilities will refer to the probability of playing red or green conditioned on the previous action, third-order probabilities will refer to the probability of playing red or green conditioned on the actions in the previous two time periods et cetera. Hence, i.i.d. behavior implies that the probabilities of a particular action being played should be the same regardless of the order, and be independent of the history of play.

Shachat and Swarthout (2004) have found that humans are able to detect deviations in first-order probabilities of play as long as they are relatively far away from the MSNE predictions. The *wfp* algorithm will not pick up on the second- and higher-order deviations from i.i.d. behavior because it does not keep track of consecutive temporal sequences of actions. In N&S's game such deviations from randomness would imply that the combinations of actions, red-red and green-green, are not played as frequently as the MSNE would prescribe, even though people may have converged to the first-order MSNE probabilities of playing red 60% of the time.

3.1 Modification of learning rules to include cases of pattern recognition

Fortunately, the *wfp* rule can be modified to allow for the detection of patterns in opponents' play. Fictitious play can be generalized to what shall be referred to as n-period (weighted) fictitious play (*fpn*) where n is an integer greater than or equal to one and refers to the complexity and depth of pattern detection[8]. For example, single-period fictitious play or *fp1*, is equivalent to the *wfp* model discussed so far because it looks at occurrences of single period actions, and therefore sequence learning or pattern detection is assumed not to exist. *Fp2* is more advanced because it observes how many times two temporally consecutive sequences of actions have been played in the past instead of only one-period actions, and then conditions the probability of an action being played on the previous period action chosen by the opponent. In this case it is assumed that a player is making use of second-order probability information[9]. For example, suppose our opponent's play is r-r-g-g-r-r-r[10],

[8] This proposed learning rule is a specific case of a class of learning rules that Fudenberg and Levine (1998) refer to as conditional fictitious play. In conditional fictitious play a set of conditional probabilities of opponents' play given the history of all players' play is used to form beliefs about an opponents' future behavior.

[9] In general any *fpn* model uses nth order adjacent probability information.

[10] Counts are created by allowing for overlapping sequences so that each action is counted twice, once as the last action in one 2-period sequence and one as the first action in another 2-period sequence. This is because there is an inherent problem in that two very different sequences can be obtained by changing when the counting starts. Also if overlapping sequences are not used then conditioning on the previous action will be problematic. Throughout the paper the following convention will apply: the written order of sequences of actions from left to right will correspond to the time order of these actions i.e. green-red refers to the strategy where green was played first and red was played in the next period.

fp2 will evaluate how often the following sequences have turned up in past play: how many times r has been followed by r, g followed by g, r followed by g and g followed by r. In this example r-r has occurred three times, g-g once, r-g once and g-r also once.

Let the subscripts i and j denote two different players, then given actions a_j and a'_j, $I_t(a_j|a'_j)$ is an indicator function that takes a value of one if a_j was the action played at time t and a'_j was the action played at time $t-1$ and takes a value of zero otherwise[11]. Define for player i, the count of a_j at time t given action a'_j as:

$$C_i(a_j|a'_j,t) = \frac{I_{t-1}(a_j|a'_j) + \sum_{u=1}^{t-2} \gamma_i^u \cdot I_{t-u-1}(a_j|a'_j)}{1 + \sum_{u=1}^{t-2} \gamma_i^u} \tag{3}$$

The *fp2* beliefs of player i of action a_j given action a'_j are then given as:

$$fp2_i(a_j|a'_j,t) = \frac{C_i(a_j|a'_j,t)}{\sum_{a_j \in S_j} C_i(a_j|a'_j,t)} \tag{4}$$

where S_j is the discrete strategy set of player j[12].

In general, the number of frequency variables or counts an *fpn* belief model must keep track of is the number of pure actions available raised to the power of n. Also, further cognitive resources are needed to remember the last $n-1$ periods of play in order to be able to condition on these. It is clear that the cognitive requirements of these algorithms increase with n at a faster than linear rate. This means that either increasing the size of the action space or the depth of pattern recognition leads to drastically higher cognitive requirements. Hence, allowing for some heterogeneity in the cognitive abilities of agents, it is not unreasonable to expect that agents will predominantly fall into the first two belief models, *fp1* and *fp2*[13].

The higher the value of n the greater the variability in the learning rule since although each count's time series is smooth, as exemplified in equation 3, because the conditioned actions change from one period to the next switching from one count to another will lead to discrete jumps in beliefs[14]. Therefore an *fp2* belief model does have the ability of explaining more variability in elicited beliefs than an *fp1* model. As in standard weighted fictitious play, *fpn* models include a memory parameter that determines the rate of memory decay or weighting of past actions.

3.2 Relaxation of the parametric form of the belief formation rule

Apart from introducing *fp*2 as an alternative variable that could influence belief formation, this paper relaxes the parametric form of the learning rules further to allow for more generalized cases, where specific parameter values will reduce the model to the usual stricter parametric forms. In

[11] At $t = 1$ the indicator function takes the value of zero for all actions a_j, since a time period $t = 0$ does not exist and therefore actions are not observable.

[12] This definition assumes that the denominator is not zero i.e. that the action a'_j has been played at least once in the past. In cases where a'_j has not been observed beliefs are assumed to be given by a uniform distribution over $a_j \in S_j$.

[13] If learning is explicit then these two rules can be defended on the basis that the number of items that must be held in short term memory is close to the psychology literature dictum of 7 plus/minus 2 items, Miller (1956).

[14] Unless if the opponent's play is purely i.i.d. up to order n, since then there will be no discernible, systematic change in beliefs as play will on average be independent of conditioning.

particular, the parametric form should include a constant because the standard *wfp* model has the following problem when estimating γ. N&S found the $\hat{\gamma}$ estimates to be very close to a value of one and they mention that this is probably due to the fact the fictitious play model is simply attempting to obtain the best fit by passing the best smooth curve through the strongly fluctuating elicited beliefs time series. Indeed, the *wfp* rule becomes smoother as γ increases which is why the lack of a constant strongly biases the value of $\hat{\gamma}$ towards one. The fact that the best fit offerred by the *wfp* model is simply to pass a smooth curve through the data implies that fictitious play models are not able to explain the variability in elicited beliefs. To allow for greater flexibility, it will be assumed that people can react to *fpn* learning rules with different levels of sensitivity. In more detail, elicited beliefs will be modeled by including a constant, α, and a coefficient of sensitivity, β, to *fpn* beliefs so that pure *fpn* belief formation is a special case. A detailed discussion of the exact parametric form of the models and justification for arriving at them are provided in Section 4.2.1.

4 Methodology

4.1 Properties of the dataset and implications for modeling

The proposed *gfpn* models are fractional response models since the dependent variable is necessarily a fraction between zero and one. Using a standard OLS linear regression to model fractional data leads to the following problems.

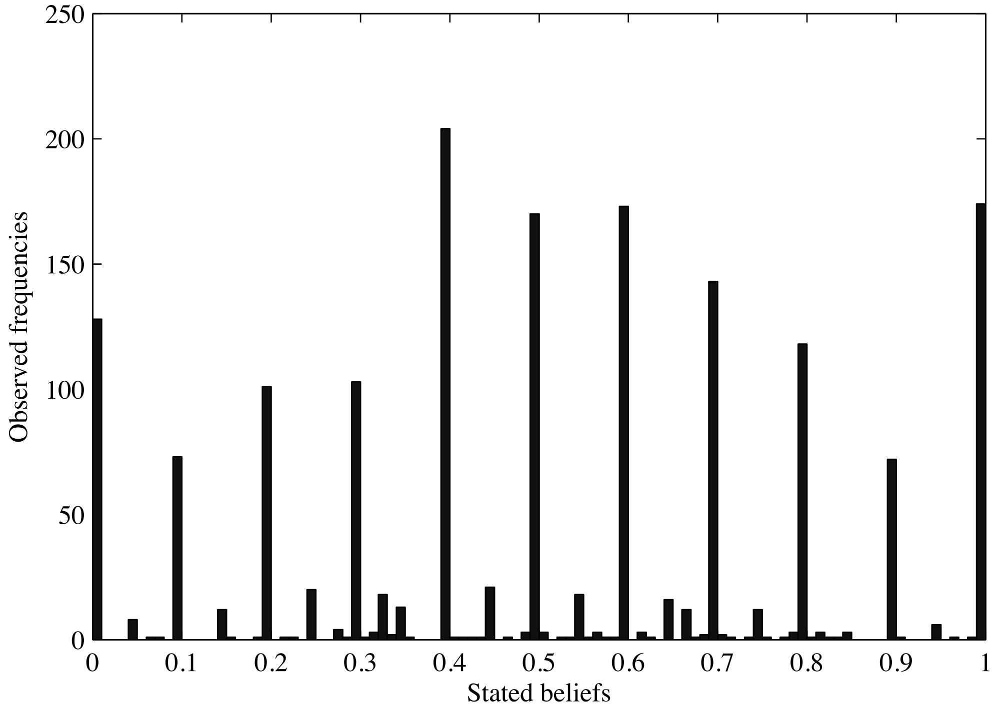
1. Unbounded predictions: Depending on the estimated coefficients of an OLS procedure the model may make impossible predictions i.e. predictions larger than one or smaller than zero, as the model's predictions are not bound in any way.
2. Errors are not normally distributed: The violation of normality arises from two sources. Firstly, since the dependent variable is bounded the error distribution is also necessarily truncated or bounded. Secondly, for predictions near the bounds of zero and one the error distribution will be highly asymmetric as errors are more likely to occur in the direction of the interior of the bounded interval. For example, for a prediction of 0.9, the errors are more likely to be negative than positive because of the truncation at one.
3. Errors are not homoskedastic: Given the bounded nature of the dependent variable there is no reason to believe that the variance at the bounds of stated beliefs will be equivalent to the variance at interior values[15].

Figure 1 displays a histogram of the dependent variable, stated beliefs. This distribution exhibits the following non-standard features.

1. Semi-continuous data: Most observations are clustered on multiples of 0.1, then on multiples of 0.05 and finally there are a few observations spread out on other values.
2. The distribution is trimodal with one peak at the interior of the interval and two peaks at the bounds. Examining the histograms of stated beliefs for each player it becomes clear that this trimodality occurs because of the existence of two main types of individual stated belief distributions. Some players exhibit a unimodal distribution with the mode somewhere in the

[15] A plausible assumption used often in the literature is that the variance of the error distribution will be greater at the midpoint of the dependent variable's bounded interval and will fall as the bounds are approached.

Figure 1 Histogram of stated belief data



interior whilst other players exhibit two modes occurring at the bounds (with or without another interior mode).

Concluding, a model must be able to reproduce these two distinct features of the distribution of stated beliefs whilst simultaneously correcting for the aforementioned three problems of standard OLS estimation. If possible, it would be conducive if the model reduced to the N&S models for special values of parameters, so that an easy comparison can be made as to the frugality or futility of the extensions proposed in this paper[16].

4.2 Econometric techniques

The econometrics literature discusses possible solutions for the three estimation problems of OLS in fitting fractional response variables, the main examples of which are discussed immediately below. However, given the irregularity of the data and the models to be estimated these proposed solutions will be shown to be problematic and an alternative method shall be proposed.

One oft used solution is to assume that the expectation of the dependent variable conditional on the data is described by a beta distribution. This approach deals with all three problems in the following ways. The problem of impossible predictions is solved as the beta distribution is necessarily bounded between zero and one (non-inclusive). Also, the beta distribution can be asymmetric for

[16] Clearly, this should not be pursued at the expense of a reasonable model structure. However, in the event that there is no other reason to choose between a number of different formulations then it is desirable to prefer the model which is better suited to carry out the desired comparisons or hypothesis tests.

certain values of its parameters and it implicitly models heteroskedasticity in a reasonable way, as the variance will be highest at a proportion of 0.5. However, the approach of assuming the dependent variable is a beta distribution is abandoned because it cannot model the extreme observations of zero and one which in this case are an integral part of the subjects' responses.

Another proposed model is the fractional logit response model, put forth by Papke and Wooldridge (1996), which is essentially a special case of the generalized linear model (GLM) employing a logit link so that the predicted values necessarily lie between zero and one. In contrast to assuming a beta distribution as mentioned above, this model is able to deal with dependent variable values of zero and one and can be estimated using maximum likelihood techniques. Standard errors are estimated using a proposed technique that is asymptotically robust to heteroskedasticity. This technique however is not applicable because there exists a non-linearity due to the γ parameter in the weighted fictitious play variables, *fp1* and *fp2*, which invalidates the heteroskedasticity-robust standard error procedure.

Due to these problems with the standard methods of dealing with fractional response variables, the following solutions are proposed as regards model structure, optimization and bootstrapping techniques.

4.2.1 Model structure The distribution is neither clearly discrete nor clearly continuous which begs the question of how to model it. One solution is a model positing discreteness such as an ordered probit model in which case the observations could be discretized by rounding to the nearest multiple of 0.1 or another solution is simply to assume a continuous distribution. The latter is more desirable in the sense that although most subjects' responses where discretized they are not necessarily restricted to be. Also, it is quite likely that the underlying process is continuous even though humans out of convenience may report probabilities as discretized values.

The *fpn* rules imply that people consciously think about using an *fpn* rule and then do so by gathering information, storing it and processing it as necessary to emulate the *fpn* algorithm i.e. it is a form of serial information processing. However, from the sequence learning literature discussed before it can be inferred that a significant part of the information processing in this case is implicit (automatically performed by parallel computation) without strict and explicit knowledge of the player. For example, although people may be reacting to past frequencies of play and basing their current play on this, their information processing may be imperfect and may not necessarily be as accurate or structured as an *fpn* algorithm implies. Players may create approximate counts of *fpn* frequencies and react to them in a heuristic or less than optimal manner. To allow for these possibilities, the following more flexible parametric forms were chosen which include a constant, α, and a coefficient of sensitivity, β, to the *fpn* variables:

$$gfp1_i(a_j, t) = \Xi[\alpha_i + \beta_i \cdot fp1_i(a_j, t, \gamma_i)] \tag{5}$$

$$gfp2_i(a_j|a'_j, t) = \Xi[\alpha_i + \beta_i \cdot fp2_i(a_j|a'_j, t, \gamma_i))] \tag{6}$$

A link function $\Xi : \mathbb{R} \to [0, 1]$ is necessary to bound the range of the model allowing it to be interpreted as a probability. Two candidates for Ξ are given in Equations 7 and 8.

$$\Xi_t(x) = min[max[x, 0], 1] \tag{7}$$

$$\Xi_l(x) = \frac{e^x}{1+e^x} \tag{8}$$

The simplest way of modeling stated beliefs and allowing for the possibility of modes occurring at the bounds is to use a function that can return values greater than zero and one, and then truncate these using simple minimum and maximum functions, as implemented in Ξ_t. Another solution is the logit function employed in Ξ_l, which is often used as a link function in Generalized Linear Models with binary outcomes.

Section 5.1 compares models estimated using both link functions and finds that the Ξ_t link function fits the data better. Henceforth, throughout this paper equations 5 and 6 will be estimated by employing Ξ_t as their link function. Apart from fitting the data better than Ξ_l, Ξ_t also has the advantage that the *wfp* model is nested within these proposed generalized models and will reduce to it if α equals zero and β equals one. Therefore there is nothing in this generalized model that imposes heterogeneity of the constant or sensitivity coefficient, or that imposes them to be different from the *wfp* models in N&S, thereby allowing for easy comparison.

The constraints on the estimated parameters imposed on the problem of estimating models are as follows[17]:

$$-1 \leq \hat{\alpha} \leq 1 \quad -2 \leq \hat{\beta} \leq 2 \quad 0 \leq \hat{\gamma} \leq 1.25 \tag{9}$$

The constraint on γ is binding at zero since there is no interpretation of a negative value, however the upper bound was set at 1.25. A value greater than one has the interpretation that a player is putting more weight on past observations than recent observations. This may occur if something akin to imprinting occurs where first impressions have a large influence on players or if players for some reason pay less attention to history as time progresses, perhaps due to fatigue.

4.2.2 Proposed measure of performance In estimating their models N&S chose to use a non-linear estimation technique which minimized the Mean Square Deviations (MSD) between stated beliefs and the predicted beliefs whereas this study opts for the use of the Mean Absolute Deviations (MAD) instead, on the following grounds.

In the data there is a large number of stated beliefs equal to the extreme values of zero and one, or in their vicinity. Hence, minimizing MSD gives the estimation technique a bias towards preferring learning rule parameters that do not lead to extreme predictions i.e. it will choose parameters that tend to smooth beliefs to a large degree. Such a bias would therefore cause memory parameter estimates to be inflated as a longer memory process serves to smooth beliefs since each additional information has a smaller effect on beliefs. Mitropoulos (2001) finds that it is generally the case that MSD minimization tends to select learning rules that make predictions closer to a uniform distribution rather than a distribution of predictions near to the bounds. In response to this problem the MAD is proposed as the error measure, since it does not excessively penalize larger absolute errors.

The MAD has other important properties that make it more suitable than MSD for this particular application. For example, Gorard (2005) provides evidence for preferring MAD over MSD particularly on the grounds of efficiency. Although Fisher (1920) defended the use of MSD over MAD by arguing that the former was more efficient than the latter, the assumptions he made were extremely strong -

[17] Imposing these bounds significantly reduces the computation time required for model estimation.

he assumed normality and no observation or measurement error. Huber (1981) did away with these strict assumptions and found that MAD is in fact more efficient than MSD whenever the percentage of error points in the total observations is greater than 0.2%[18]. Wilson (1978) surmises that MAD is dramatically more efficient than MSD in the presence of outliers contaminating the dataset. The superiority of MAD over MSD for non-normal distributions was pointed out by Fisher himself and has been verified through the use of Monte Carlo techniques. This is extremely important for this study as error distributions of fitted stated beliefs will clearly be non-normal as argued earlier. In addition to this, the error distributions of the estimated models will not be evenly distributed around the function, i.e. will be asymmetric, in which case Matheson (1990) concludes again that MAD should be preferred over MSD.

4.2.3 Optimization technique The models to be estimated are intrinsically nonlinear in the memory parameter and therefore standard regression techniques cannot be used as a closed form solution for minimizing the MAD does not exist. The optimization-search technique used to solve the models is a hybrid procedure which uses genetic algorithms to pick a good starting point for the Nelder and Mead (1965) simplex method, chosen for its ability to function well with non-smooth or discontinuous functions. Standard optimization methods that are based on gradient descent techniques could become stuck in local minima easily because the MAD function is not necessarily smooth. Details of the optimization techniques employed and a discussion of their advantages over more commonly used techniques is provided in Appendix B. The necessary code to run these routines was written in Matlab (2007).

4.2.4 Bootstrapping technique N&S did not estimate any confidence intervals for the estimated parameters and therefore it was not possible test for the statistical significance of the memory parameter estimates. The non-linear nature of these learning rules automatically poses a difficulty in estimating confidence intervals for coefficients, especially because error terms are not Gaussian. This will be dealt with by using a non-parametric bootstrapping method to approximate the distribution of the estimated coefficients. Inspection of the stated belief data on an individual basis reveals that the distribution of stated beliefs can be very different from player to player i.e. the error distributions exhibited significant between-subjects heterogeneity. Also, the error distributions showed within-subjects heterogeneity as the shape of the error distributions are conditional on the predicted stated beliefs - a perfectly reasonable observation given the bounded nature of the dependent variable. Hence, any bootstrapping procedure used must sample from the same individual and for the same predicted value in order to account for both of these types of heterogeneity. Also, if possible the bootstrapped stated beliefs should follow the features of the original data in particular as regards to the predominance of stated beliefs given as multiples of 0.1.

The random-x resampling procedure (Davison and Hinkley, 1997) is clearly not appropriate because of the dependence of the *fpn* variables on their previous values which would be disrupted by such a procedure. The commonly used fixed-x resampling procedure of drawing errors from the error distribution and adding these sample errors to the predicted value of the dependent variable, $\widehat{sb_t}$, was not adopted for the following reasons. Such a procedure may lead to values greater than one or less than zero and will lead to a smoother distribution of bootstrapped stated beliefs as they will not lie on multiples of 0.1.

[18] If the percentage of error points is 5% then MAD is twice as efficient as MSD.

Instead of sampling from the error distribution and adding these error samples to the fitted stated beliefs to obtain the bootstrapped stated belief, a similar approach which does not have the above problems is to directly sample from the stated beliefs conditional on the value of the fitted stated beliefs. More specifically, each model will be estimated once for each player with the original stated beliefs data, sb_t, and the correspondence of the stated beliefs and the predicted stated beliefs after rounding them to the nearest multiple of 0.1, $\widehat{sb_t}$, will be recorded (i.e. for each observation a value of sb_t will correspond to a value of $\widehat{sb_t}$). Bootstrapping for each observation will proceed in the following manner. The stated beliefs to be used in each bootstrap regression will be randomly sampled from all the values of the original stated beliefs that correspond to each predicted stated belief, $\widehat{sb_t}$. For example, if for a particular observation the predicted stated belief in the original model is 0.6, then the bootstrapped stated belief will be drawn with replacement from the values of stated beliefs where the corresponding predicted stated belief was equal to 0.6. This process will be run independently for each player in order to allow for heterogeneity between subjects. For each individual, 1,000 bootstrapped models will be estimated and from the resulting distribution of coefficient values the 2.5% and 97.5% percentiles are used as the bounds of the 5% confidence interval. In this case, optimizing this model each time using Genetic Algorithms to elicit starting points for the Nelder and Mead (1965) simplex method is too inefficient in terms of computation time. Hence, for each of the 1,000 repetitions the starting values for the Nelder and Mead (1965) simplex method will be the original parameter estimates obtained from the procedure on the original, non-bootstrapped stated beliefs series. The reasoning behind this is that the new bootstrapped parameter estimates should be in the vicinity of the original estimates and therefore the original parameter estimates should make for efficient starting values.

5 Results

5.1 Comparison of link functions

Non-linear regressions of the models in equations 5 and 6 were estimated with the two different link functions discussed in Section 3.2 in an effort to determine which is more suitable. The MAD averaged across all players was calculated for both Ξ_t and Ξ_l and both *gfp1* and *gfp2* models. A comparison of these two measures of fit leads to the conclusion that Ξ_t provides consistently better fit than Ξ_l. For the *gfp1* model the average MAD of all players were 10.43 and 10.77 for Ξ_t and Ξ_l respectively, and for the *gfp2* model 10.38 and 10.56 respectively. Since for both cases closer fit is achieved with Ξ_t, and it also conveniently reduces to standard weighted fictitious play, a property which Ξ_l does not posses, Ξ_t is chosen as the link function for all the estimated models.

5.2 Comparison of the standard weighted fictitious play model to gfp1 and gfp2 models

Before proceeding further it is important to ascertain whether the generalized model proposed in this paper is justified compared to the simpler *wfp* model used in N&S. To testify to this, a comparison can be made between the MSD of the *gfpn* models proposed here and the original *wfp* models of N&S[19]. The total sum of squared errors from the N&S formulation is 145.7, whereas the total sum

[19] Note that this comparison puts the *gfpn* models at a disadvantage because the N&S model was solved by minimizing MSD whereas the *gfpn* models were estimated by minimizing MAD. Hence, the MSD for *gfp1* and *gfp2* could be reduced further in practice.

of squared errors of the *gfp1* model is 114.1 and that of the *gfp2* model, 108.5. The total sum of squared errors of the standard model is 27.7 % higher than that of *gfp1* and 34.3% higher than the *gfp2* model[20]. Hence, it appears that the use of the *gfpn* models is justified, even though there is an increase in the number of free parameters, from one (just the memory decay parameter) to the three free parameters of the generalized models.

Also, a comparison can be made between the models and the MSNE prediction which would imply that a player's belief that his opponent will play red should be 0.6. The MAD averaged across subjects for MSNE beliefs is 14.95 in contrast to those for *gfp1* and *gfp2* models for all players which are 10.43 and 10.38 respectively, a reduction in the MAD by roughly a third. Therefore, the hypothesis of MSNE beliefs is clearly refuted by the elicited beliefs data.

5.3 Estimates of the sensitivity coefficient, $\hat{\beta}$, and significance of the fp1 and fp2 variables

One of the most important contributions in this reanalysis is the ability to ascertain whether *fp1* and *fp2* variables are statistically significant or whether a simple constant is equally effective at predicting elicited beliefs. This was achieved by constructing 95% confidence intervals from the bootstrap and then observing whether the sensitivity coefficient is significantly different from zero.

The average value of $\hat{\beta}$ for the *gfp1* models is 0.505 (0.59 for $\hat{\beta}$ values significantly different from zero) and 0.503 (0.82 for $\hat{\beta}$ values significantly different from zero) for *gfp2* models. The *fp1* variable is significant in 17 of the 28 players, whereas the *fp2* variable is significant for 13 players, with these results overlapping in 11 of these cases, indicating that the *gfp1* and *gfp2* models are directly competing models of subjects' behavior.

A hypothesis that subjects on average are underreacting to *fpn* variables is strongly support by the evidence since fourteen estimates of β were found to be significantly less than one for the *gfp1* models and seventeen for the *gfp2* models.

Another innovation in this analysis was that $\hat{\beta}$ values are not constrained to be between -1 and 1 and therefore it is possible to model player over-reaction to the *fp1* and *fp2* variables and/or risk-seeking behavior[21] (without the ability however to differentiate between the two). In the *gfp2* models of players there are two sensitivity coefficients whose absolute value is greater than one, whereas for the *fp1* models there are seven cases.

Negative values of $\hat{\beta}$ can be associated with players who believe that for some reason there is likely to be mean reversion in opponents' play. As mentioned before this could be due to expectations of poor randomization or because of sophisticated strategies where players change their mode of play in order to gain a competitive advantage. There are three negative $\hat{\beta}$ coefficients significantly different from zero for *gfp1* models and none for the *gfp2* models.

[20] These comparisons imply homogeneity of players as they assume that either all players are *gfp1* or all players are *gfp2*. In Section 5.6 the assumption of homogeneity of players will be dropped so that players will be classified as following the *gfp1* model or *gfp2* model. Allowing for player heterogeneity, the desirability of the *gfpn* models compared to the standard weighted fictitious play model will become even more acute since for each individual the best fitting *gfpn* model will be chosen.

[21] Appendix A discusses how the quadratic scoring rule when combined with risk-seeking preferences might lead to stated beliefs near the bounds.

5.4 Estimates of the memory parameter, $\hat{\gamma}$

N&S's estimates of the memory parameter were centered on one with very little dispersion, implying that individuals weighted all past information equally. These results however do not stand up in this reanalysis as can be seen in Tables 2 and 3. For the *gfp1* model the average estimated $\hat{\gamma}$ is 0.792 and for the *gfp2* model it equals 0.882. More importantly, the average values of $\hat{\gamma}$ conditional upon the sensitivity coefficient, $\hat{\beta}$, being significantly different from zero are 0.548 and 0.794 respectively. In the seventeen cases where the *gfp1* model exhibited $\hat{\beta}$ significantly different from zero, the values of $\hat{\gamma}$ for six players are equal to zero. From this it can be inferred that the *gfp1* model is often capturing a special case of weighted fictitious play behavior, namely generalized Cournot adjustment[22].

An interesting result is the difference between the estimates of γ for *gfp1* and *gfp2* models. Although *gfp2* is cognitively more demanding than *gfp1*, players exhibited less memory decay. This is not surprising as the *gfp2* model essentially needs a high valued $\hat{\gamma}$ as it requires more memory depth in order to be able to detect patterns. In fact, the *gfp1* models exhibited 8 cases where $\hat{\gamma}$ is equal to zero, 6 of which are coupled with a sensitivity coefficient, $\hat{\beta}$, significantly different from zero. This is in stark contrast to the *gfp2* models where this occurs in three cases, only one of which is associated with a sensitivity coefficient significantly different from zero.

5.5 Estimates of the constant/anchoring coefficient, $\hat{\alpha}$

If players were using *wfp* beliefs then the values of $\hat{\alpha}$ should be close to zero. For the *gfp1* models $\hat{\alpha}$ is significantly different from zero at the 5% level in 15 out of the 28 players, whereas for the *gfp2* models this occurs for 19 out of 28 players. The average values of $\hat{\alpha}$ for *gfp1* and *gfp2* models are 0.23 and 0.292 respectively. Individual values of $\hat{\alpha}$ were not restricted to be greater than or equal to zero and indeed in some cases were estimated to be less than zero. Five of the *gfp1* estimated models had negative estimates of α, however only one of these was significantly different from zero at the 5% level. Out of the *gfp2* models only one $\hat{\alpha}$ was estimated to be less than zero, and this estimate also happened to be statistically significant from zero.

In conclusion, there is significant evidence of individuals using anchoring or having some prior over opponents' beliefs, and adjusting the predictions from this anchor by incorporating *fpn* variables into their final decision. Also, the high predominance of $\hat{\alpha}$ values significantly different from zero strengthens the argument in Section 3.2 that not including a constant in the learning model leads to an adverse effect on the estimates of γ.

5.6 Classification of players according to belief formation models

A central aim of this study was to determine whether any players used pattern recognition learning algorithms and if so, to attempt to estimate how prevalent such learning rules are in the N&S subject pool. In this section, players will be classified either as *gfp1* players, *gfp2* players, *gfp1*/*gfp2* players or *non-gfpn* players, as there may be some other learning rule not examined in this paper that

[22] It is referred to as generalized in the sense that $\hat{\alpha}$ is not necessarily equal to zero and $\hat{\beta}$ is not necessarily equal to one as would be the case in standard Cournot adjustment. In this particular game, Cournot adjustment is equivalent to both a reinforcement learning model and fictitious play model which only look at the previous period, and to the win-stay/lose-shift heuristic that will be discussed later.

Table 2 *Gfp1* non-linear regression results and upper and lower parameter bootstrapped estimates of 95% confidence intervals

Player	MAD	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}^{lower}_{95\%}$	$\hat{\alpha}^{upper}_{95\%}$	$\hat{\beta}^{lower}_{95\%}$	$\hat{\beta}^{upper}_{95\%}$	$\hat{\gamma}^{lower}_{95\%}$	$\hat{\gamma}^{upper}_{95\%}$
1	9.20	-0.47	2.000*	0.786*^	-0.521	0.449	0.554	2	0.597	0.975
2	4.91	0.397*	0.009^	1.151*	0.333	0.7	-0.852	0.3	0.523	1.25
3	20.24	-0.521	2	0.871*	-0.875	0.861	-0.783	2	0.595	1.139
4	11.31	0.600*	-0.695*^	0.967*	0.65	1	-2	-0.57	0.697	1.25
5	7.65	0.900*	-0.800*^	0^	0.807	1	-0.9	-0.622	0.000	0.127
6	15.46	0.056	1.249*^	1.102*^	-0.17	0.028	2	2	1.055	1.25
7	9.50	0.123	0.584*^	0.535*^	-0.047	0.296	0.346	0.848	0.348	0.757
8	11.35	0.527	0.039^	0.653	-0.015	0.999	-0.89	0.898	0.000	1.248
9	9.60	0.369*	0.343*^	0.364*^	0.113	0.399	0.218	0.894	0.078	0.834
10	12.90	0.001	0.613	0.885*	-0.84	1	-1.397	2	0.245	1.25
11	8.10	1.000*	-1.004*^	0^	1	1	-1.01	-0.95	0.000	0.01
12	12.96	0.336	0.640*	0.753	-0.287	0.6	0.102	1.962	0.000	1.144
13	7.07	0.247	0.408*^	1.247*	-0.061	0.327	0.323	0.761	0.86	1.25
14	7.77	-0.073	1.065	1.25*	-0.467	0.999	-0.613	1.661	0.418	1.25
15	14.05	0.332*	0.268^	0^	0.1	0.6	-0.000	0.746	0.000	0.686
16	13.91	0.250*	0.122^	0.894^	0.25	0.5	-0.511	0.386	0.101	1.25
17	16.98	0.2	0.600*	0^	-0.585	0.4	0.203	1.464	0.000	0.396
18	11.50	-0.633	1.533*	1.237*^	-0.988	0.127	0.601	2	1.02	1.25
19	2.62	0.450*	0.150*^	0^	0.449	0.499	0.101	0.152	0.000	0.098
20	6.60	0.400*	0.2^	0^	0.398	0.6	-0.000	0.204	0.000	0.114
21	2.90	0.500*	0^	1.099*^	0.5	0.5	-0.000	0	1.055	1.22
22	8.00	0.500*	0.506*	0^	0.389	0.5	0.505	2	0.000	0.251
23	5.48	0.380*	0.241^	0.483^	0.273	0.6	-0.2	0.486	0.000	0.962
24	8.40	0.500*	0.1*^	0^	0.397	0.599	0	0.294	0.000	0.551
25	10.00	0.76	-0.46	0.969	-0.916	1	-0.856	2	0.000	1.25
26	11.80	0.293*	0.416*^	0.277^	0.175	0.599	0.101	0.62	0.000	0.646
27	15.36	-0.161	2.000*^	1.066*	-0.197	0.18	1.997	2	0.669	1.093
28	16.51	-0.834*	2.000*^	0.984*	-0.893	-0.659	1.999	2	0.866	1.034
Average	10.43	0.23	0.505	0.628						

Parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ are superscripted by * if significantly different from zero (5% level)
Parameters $\hat{\beta}$, $\hat{\gamma}$ are superscripted by ^ if significantly different from one (5% level)

Table 3 *Gfp2* non-linear regression results and upper and lower parameter bootstrapped estimates of 95% confidence intervals

Player	MAD	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}^{lower}_{95\%}$	$\hat{\alpha}^{upper}_{95\%}$	$\hat{\beta}^{lower}_{95\%}$	$\hat{\beta}^{upper}_{95\%}$	$\hat{\gamma}^{lower}_{95\%}$	$\hat{\gamma}^{upper}_{95\%}$
1	11.97	0.375*	0.581^	1.22*	0.225	0.800	-0.000	0.775	0.891	1.250
2	4.26	0.509*	-0.109^	0.138^	0.400	0.600	-0.200	0.000	0.000	0.335
3	19.35	0.002	0.989*	0.839	-0.760	0.453	0.300	2.000	0.000	1.031
4	11.50	0.400*	0.2^	0^	0.300	0.600	-0.270	0.400	0.000	0.695
5	4.86	0.100*	0.722*^	0.725*^	0.095	0.100	0.700	0.803	0.769	0.807
6	15.29	0.200*	0.600*	1.056*	0.008	0.314	0.213	1.765	0.779	1.250
7	9.34	0.1	0.609*^	0.761*^	-0.079	0.292	0.337	0.866	0.503	0.894
8	11.17	0.400*	0.238^	1.137*	0.263	1.000	-1.117	0.509	0.778	1.250
9	6.25	0.270*	0.479*^	0.751*^	0.184	0.300	0.400	0.614	0.081	0.864
10	12.54	0.150*	0.35^	0^	0.097	0.400	-0.100	0.500	0.000	0.718
11	2.80	0.002*	0.999*	0.72*^	-0.330	-0.213	1.832	2.000	0.885	0.949
12	13.32	0.536*	0.164^	0.985	0.105	0.832	-0.140	0.596	0.000	1.250
13	6.72	0.332*	0.318^	1.093*	0.259	0.500	-0.000	0.445	0.775	1.250
14	7.94	0.490*	0.159^	1.014*	0.129	0.998	-0.580	0.586	0.365	1.250
15	14.74	0.233	0.267	1.118*	-0.729	1.000	-0.900	1.995	0.710	1.250
16	14.05	0.311	0.025^	1.067	0.000	0.700	-0.464	0.466	0.000	1.250
17	19.87	0.084	0.934	1.103*	-0.709	0.700	-0.808	2.000	0.939	1.250
18	10.52	0.002	0.898*	1.081*	-0.800	0.253	0.447	2.000	0.944	1.250
19	3.53	0.445*	0.155*^	0.482^	0.409	0.550	0.000	0.196	0.000	0.850
20	7.62	0.592*	0.008^	0.830	0.300	0.763	-0.455	0.373	0.000	1.250
21	2.99	0.476*	0.035^	0.98*	0.500	0.500	-0.000	0.000	0.967	1.083
22	12.14	0.576	2.000*	0.57*^	-0.081	0.879	1.000	2.000	0.126	0.912
23	4.30	0.400*	0.2*	0^	0.400	0.600	0.000	0.200	0.000	0.253
24	8.68	0.395	0.221^	0.969*	-0.141	0.600	-0.000	0.857	0.255	1.200
25	10.21	0.494*	0.007^	0.687	0.150	0.800	-0.500	0.576	0.000	1.250
26	12.54	0.155	0.766*	1.25*	-0.282	0.515	0.265	1.394	1.095	1.250
27	13.57	0.200*	1.454*^	1.075*	0.200	0.300	1.241	1.454	1.049	1.095
28	18.55	-0.052*	0.807*^	0.985*	-0.999	-0.415	1.598	2.000	0.778	1.232
Average	10.38	0.292	0.503	0.808						

Parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ are superscripted by * if significantly different from zero (5% level)
Parameters $\hat{\beta}$, $\hat{\gamma}$ are superscripted by ^ if significantly different from one (5% level)

better describes them. An obvious way to differentiate the *gfpn* players from the *non-gfpn* players is to examine whether the estimated sensitivity coefficients, $\hat{\beta}$, of the *fp1* and *fp2* variables were statistically different from zero[23]. However, since the *fp1* and *fp2* variables could not be included in the same equation due to their high degree of correlation, such a method would not allow us to make direct comparisons between *gfp1* and *gfp2* players which is an integral part of this study.

A solution to this problem can be based on the bootstrapped equations estimated earlier which provide bootstrapped distributions of the MAD for the *gfp1* and *gfp2* models. The baseline model, *non-gfpn*, will be a simple model with only a constant and error term i.e. it will simply fit a time-invariant estimate of beliefs implying that the variability of stated beliefs is random. In this light, the same bootstrapping procedure applied in Section 4 is used to obtain an estimated distribution of the MAD for this baseline model with which to compare the postulated models.

Given the three models that are posited there are three possible pairwise comparisons that can be made between the distributions of the MAD for each player and model. The bootstrapped distributions of the MAD for each model and player were found to exhibit non-constant variance and often deviated significantly from normality assumptions. Hence, statistical testing of the difference in the MADs was accomplished through the use of bootstrapped statistics. The exact procedure is as follows. From the bootstrapped regressions performed earlier there exist 1,000 bootstrapped observations of the MAD for each player and each estimated model. From each such distribution 1,000 samples are drawn with replacement and the average of these values is calculated. The difference of these means between the pair of models under consideration is recorded and by repeating this procedure 1,000 times one obtains the distribution of the difference of the means of MAD for each player and model. It is then simple to test whether the difference in the mean of the MAD distributions is different from zero by constructing a confidence interval from the appropriate percentiles of this distribution. If the significance level of each individual test or the per-comparison error rate (PCER), the number of pairwise tests performed is n, then the family-wise error rate (FWER), assuming independence, is determined by a binomial distribution:

$$FWER = 1 - (1 - PCER)^n \tag{10}$$

Hence, if the PCER is 5%, then for the three possible pairwise comparisons the probability of making a Type I error in any one of those tests is now 14.26%. The effects of multiple comparisons can be controlled for by using the Sidak (1967) correction. According to this correction, since the family-wise error rate (FWER) for n pairwise comparisons is related to the PCER according to equation 10, it is sufficient to set the FWER at the desired level of significance instead of directly setting the PCER. Hence, all confidence intervals will be calculated at a 5% FWER, which will determine the respective Sidak-corrected PCER from equation 10. When comparing the *gfp1* and *gfp2* models to the *non-gfpn* model one-tailed tests will be employed as the *non-gfpn* model is nested within the *gfp1* and *gfp2* models and therefore the fit will necessarily be better for these two models. Hence, the appropriate alternative hypothesis is that the errors from these models are less than those of the *non-gfpn* models. However, in the pairwise comparison of the *gfp1* and *gfp2* model there is no a priori reason to expect that one will be better than the other and therefore two-tailed tests will be performed.

[23] Such a comparison implies that *non-gfpn* models are essentially models of players with a fixed average belief and random fluctuations around it.

Calculating all pairwise comparisons will lead to a ranking of the three models in terms of the statistical significance of differences in their MADs. However, this ranking is not necessarily transitive. If in all three tests we reject the null hypothesis then the ranking of the three models based on this will necessarily be transitive. If for at least one of the three tests one of the results is that the null hypothesis is not rejected then it is possible for the rankings to be intransitive. Let $a \succ b$ denote that model a has a statistically significant lower MAD than model b, and that $a \sim b$ denotes that a and b do not have statistically significant different MADs so that either model is valid in explaining subjects' behavior. One possibility is the following: *gfp1* $\succ$ *non-gfpn*, *gfp2* $\sim$ *non-gfpn* but *gfp1* $\sim$ *gfp2* which may occur if *gfp1* has slightly better fit than *gfp2* so that they are not statistically significant but where this slight advantage of *gfp1* may be enough to make it statistically different from the *non-gfpn* model. No such case however occurs in the dataset and therefore this possibility is not of concern. In the cases where *gfp1* $\sim$ *non-gfpn* and *gfp2* $\sim$ *non-gfpn* the player will be classified as a *non-gfpn* player since the additional information provided by the inclusion of the *fp1* and *fp2* variables does not significantly increase the fit of the models. This conclusion is preferred on the basis of Occam's razor, in other words assuming that subjects are using the simplest and least computationally expensive model of the three which is clearly the *non-gfpn* model. This classification is further supported by the fact that by construct the *non-gfpn* model will not outperform the other two models and therefore we will never observe that *non-gfpn* $\succ$ *gfp1* or *non-gfpn* $\succ$ *gfp2* . It is also possible that *gfp1* $\succ$ *non-gfpn*, *gfp2* $\succ$ *non-gfpn* but *gfp1* $\sim$ *gfp2* so that we may have a case where both the *gfp1* and *gfp2* models are better than the *non-gfpn* model but we cannot decide whether *gfp1* or *gfp2* is better with some measure of statistical certainty. In this case, we will lean on the side of cautiousness and will classify the model as *gfp1/gfp2*. The alternative of simply choosing the model with the lowest MAD will very often lead to erroneous results as the differences may be due to chance.

The results of these tests are given in Table 4. If the upper bounds of the first two columns are less than zero then this means that the *gfp1* and *gfp2* models are significantly better than the *non-gfpn* model. If both the lower and upper bounds of the difference in the MAD of the *gfp1* and *gfp2* models, given in the third and fourth columns, are less than zero then the *gfp2* model gives a statistically significant better fit than the *gfp1* model. The *gfp2* model best describes 35.7% of the subjects, 39.3% are best described by the *gfp1* model, 10.71% are classified as *non-gfpn* and in 14.3% of the cases we could not distinguish between the *gfp1* and *gfp2* models but could reject the assumption that these players were *non-gfpn*. Due to control of the family-wise error rates through the use of the Sidak correction these results assure us of the existence of heterogeneity and of the importance of modeling it.

The occurrence of cases where it is not possible to conclude on the basis of statistical significance whether the *gfp1* or *gfp2* model is better is not a weakness rather it is a result of the fact that in many cases the two models make very similar predictions. In particular, if an opponent's play is serially uncorrelated then there is not much for the *gfp2* algorithm to pick up over the *gfp1* algorithm apart from small deviations due to random fluctuations. It is expected that in such a case the two models will yield highly correlated belief predictions and therefore similar MADs for the following reason. Assume a player deviates from first-order probabilities but his second-order probabilities are i.i.d. given his first-order probabilities. Under such circumstances, the probabilities of playing each action conditional on the previous period action should all be approximately equal

Table 4 Lower and upper bounds of differences in MAD of each model in pairwise Sidak-corrected comparisons (5% experimentwise significance level)

	$upper^{1-tail}_{98.3\%}$ of	$upper^{1-tail}_{98.3\%}$ of	$lower^{2-tail}_{98.3\%}$ of	$upper^{2-tail}_{98.3\%}$ of		Statistics of classified model			
Pl.	*gfp1* − *non-gfpn*	*gfp2* − *non-gfpn*	*gfp2* − *gfp1*	*gfp2* − *gfp1*	Classification	MAD	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
1	-2.42	-0.39	1.77	2.16	*gfp1*	9.20	-0.47	2.000	0.786
2	0.05	-0.68	-0.81	-0.64	*gfp2*	4.26	0.509	-0.109	0.138
3	-1.11	-2.70	-1.83	-1.33	*gfp2*	19.35	0.002	0.989	0.839
4	-0.37	-0.09	0.13	0.43	*gfp1*	11.31	0.600	-0.695	0.967
5	-10.70	-11.69	-1.12	-0.79	*gfp2*	4.86	0.100	0.722	0.725
6	-0.24	-0.21	-0.17	0.22	*gfp1/gfp2*	15.46/15.29	0.056/0.200	1.249/0.600	1.102/1.056
7	-2.82	-3.11	-0.41	-0.16	*gfp2*	9.34	0.1	0.609	0.761
8	-0.20	-0.13	-0.08	0.24	*gfp1/gfp2*	11.35/11.17	0.527/0.400	0.039/0.238	0.653/1.137
9	-1.48	-4.58	-3.20	-2.98	*gfp2*	6.25	0.270	0.479	0.751
10	-0.55	-0.39	-0.03	0.31	*gfp1/gfp2*	12.90/12.54	0.001/0.150	0.613/0.35	0.885/0
11	-14.35	-17.98	-3.74	-3.33	*gfp2*	2.80	0.002	0.999	0.72
12	-0.97	-0.16	0.66	0.98	*gfp1*	12.96	0.336	0.640	0.753
13	-0.20	0.45	0.53	0.76	*gfp1*	7.07	0.247	0.408	1.247
14	0.25	0.22	-0.17	0.09	*non-gfpn*	8.05	0.6		
15	-1.23	-0.44	0.62	0.98	*gfp1*	14.05	0.332	0.268	0
16	0.03	0.03	-0.25	0.22	*non-gfpn*	14.03	0.33		
17	-3.81	-1.59	1.98	2.48	*gfp1*	16.98	0.2	0.600	0
18	-0.83	-1.87	-1.18	-0.87	*gfp2*	10.52	0.002	0.898	1.081
19	-1.31	-0.62	0.63	0.76	*gfp1*	2.62	0.450	0.150	0
20	-0.65	-0.07	0.48	0.68	*gfp1*	6.60	0.400	0.2	0
21	0.14	0.14	-0.13	0.11	*non-gfpn*	2.9	0.5		
22	-4.04	-0.70	3.05	3.64	*gfp1*	8.00	0.500	0.506	0
23	-0.28	-0.65	-0.47	-0.30	*gfp2*	4.30	0.400	0.2	0
24	-0.48	-0.41	-0.08	0.22	*gfp1/gfp2*	8.40/8.68	0.500/0.395	0.1/0.221	0/0.969
25	0.14	-0.18	-0.44	-0.18	*gfp2*	10.21	0.494	0.007	0.687
26	-1.87	-1.02	0.64	1.03	*gfp1*	11.80	0.293	0.416	0.277
27	0.35	-1.63	-2.42	-1.54	*gfp2*	13.57	0.200	1.454	1.075
28	-4.46	-4.13	0.12	0.52	*gfp1*	16.51	-0.834	2.000	0.984

Table 5 Means of estimated parameters grouped by classified models

Model	Means of estimated parameters $\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
gfp1	0.187	0.59	0.456
gfp2	0.208	0.625	0.678
non-gfpn	0.477		

to the unconditional probabilities of playing each action. Hence, a *gfp2* algorithm would perform very similarly to a *gfp1* algorithm and it would be impossible to differentiate between the two. This is an important result as it implies that if an opponent's play is i.i.d. then even if a player is using *gfp2* it may be very difficult to detect this econometrically.

5.6.1 Parameter estimates of classified subjects The last four columns of Table 4 provide the estimated parameters and MAD for each player for the models which they were classified as belonging to. Table 5 documents the mean values of these estimated parameters for each type of classified model. It is clear that there do not exist not large differences in estimates of $\hat{\alpha}$ and $\hat{\beta}$ for players classified as *gfp1* or *gfp2*. The average values of $\hat{\beta}$ for both models are clearly significantly less than one supporting the inclusion of the β parameter on the basis that humans do not exhibit full sensitivity to *fpn* variables due to cognitive constraints. However there is a significant difference in the estimated memory parameter, $\hat{\gamma}$, which is equal to 0.456 for players classified as *gfp1*, and equal to 0.678 for players classified as *gfp2*. It is also interesting to note the following statistics not included in the table. For five out of the eleven players classified as *gfp1*, $\hat{\gamma}$ was equal to zero (in another two cases it was not significantly different from zero), implying that they are behaving according to a generalized Cournot adjustment process. The result for *gfp2* players stands in stark contrast as from the ten *gfp2* players only in one case was $\hat{\gamma}$ equal to zero.

For the three players classified as *non-gfpn* it is interesting to investigate whether their beliefs are MSNE beliefs in which case $\hat{\alpha}$ would be equal to 0.6. From the bootstrapped confidence intervals it was found that only for Player 14 was $\hat{\beta}$ not significantly different from the MSNE beliefs.

5.6.2 On the heterogeneity of behavior The above results are in accord with other research in experimental game theory that supports findings of heterogeneity and the importance of allowing for this heterogeneity when modeling behavior. The *gfp2* model appears to have considerable merit as roughly the same number of players are classified as *gfp2* and *gfp1*, ten versus eleven. The importance of the *gfp2* model becomes more evident if one takes into account that most *gfp1* players are essentially following a generalized Cournot adjustment learning model and therefore are not keeping track of the complete history of first-order play, focusing only on the immediately prior period. Also, there are 4 players for whom it was not possible to decide whether they followed *gfp1* or *gfp2* and who therefore potentially are *gfp2* players.

Another important result is that some players could not be assigned to either model, and that a simple constant was almost just as effective in predicting elicited beliefs. Perhaps these players' beliefs are dependent on some exogenous noisy signal and therefore are best modeled by random beliefs, or it is possible that some other model that was not specified is able to predict their behavior. This is why such subjects are referred to as *non-gfpn* players rather than random belief players as

Table 6 Belief model classification of pairs of players, number of pairs

	non-gfpn	*gfp1*	*gfp2*	*gfp1/gfp2*
non-gfpn	1 (0.16)	0 (1.18)	1 (1.07)	0 (0.43)
gfp1	-	3 (2.16)	3 (3.93)	2 (1.57)
gfp2	-	-	2 (1.79)	2 (1.43)
gfp1/gfp2	-	-	-	0 (0.29)

Expected no. of pairs, assuming independence, in parentheses

there is always the possibility that there exists some other model not examined in this study that may explain their behavior.

Despite the discovery of subject heterogeneity as to whether pattern detection was employed or not it is difficult to conclude whether this heterogeneity can be attributed to within-subjects heterogeneity or between-subjects heterogeneity. This distinction has important consequences, since in the case of between-subjects heterogeneity it is implicitly assumed that each subject does not have access to the use of pattern detecting models, perhaps due to cognitive bounds. In this application, this would suggest that some players never use a *gfp2* model either because they do not have the cognitive ability or perhaps for some other reason they are not privy to this rule. Within-subjects heterogeneity, on the other hand, implies that each player has the ability to use a *gfp1* or *gfp2* learning model and determines which strategy to use based on the opponent's behavior. The researcher will normally only observe one realization of the two models being used, thereby making it difficult to determine whether the model not observed is in fact accessible to the subject. However, this does not mean that the subject cannot use a *gfp2* model but simply did not choose to do so. In this study, it is not unreasonable to assume that whether a *gfp1* or *gfp2* learning rule is adopted by a player depends on the behavior of their opponent. For example, if one's opponent does not exhibit second-order deviations from i.i.d. behavior then a player may stick to the *gfp1* rule which has lower cognitive and information processing costs, whereas if an opponent does exhibit patterns in his play then the player may use a *gfp2* rule instead.

There is in fact some evidence from the data in Table 7 that this may be true. Let the temporal structure and properties of chosen actions be referred to as an action profile henceforth. Table 7 matches each subject's action profile with the classified belief model of that subject's opponent. MSNE(1) and MSNE(2) are abbreviations denoting that behavior was not found to be significantly different from first-order MSNE play and second-order MSNE play respectively, whereas a preceding $\sim$ denotes that behavior was found to be significantly different from the MSNE prescriptions[24]. Such a comparison sheds light upon the question of whether subjects tried to take advantage of non-MSNE behavior in opponents when it existed. The ratio of the probability of a player being classified as *gfp2* to being classified as *gfp1* when the opponent is either $\sim$MSNE(1)/$\sim$MSNE(2) or MSNE(1)/$\sim$MSNE(2) (i.e. does not conform to the MSNE prescription for second-order play) is 5:3, whereas the same ratio when the opponent is MSNE(1)/MSNE(2) decreases to 5:8. This result is supporting evidence that the subjects were more likely to be classified as *gfp2* rather than *gfp1* when their opponent exhibited statistically significant deviations from MSNE(2), and therefore indirectly supports the notion that there may exist significant within-subjects heterogeneity in the subject pool. A two-tailed Fisher exact test on these proportions generates a p-value of 0.387, so that the

[24] These results are based on χ^2 tests, details of which are given in Section 5.7.

Table 7 Player matching based on subjects' action profiles and opponents' belief model classification

	∼MSNE(1)/∼MSNE(2)	MSNE(1)/∼MSNE(2)	MSNE(1)/MSNE(2)
non-gfpn	0	1	2
gfp1	3	0	8
gfp2	3	2	5
gfp1/gfp2	3	0	1

Table 8 Distribution of actual and MSNE prescribed play

	% of combinations of red and green action choices					
	r	g	g-g	g-r	r-g	r-r
MSNE predictions	60	40	16	24	24	36
All players, 60 rounds	52.44	47.56	21.43	26.33	26.03	26.21
Row Players, 60 rounds	51.43	48.57	23.85	24.94	24.58	26.63
Column players, 60 rounds	53.45	46.55	19.01	27.72	27.48	25.79
All players, last 30 rounds	51.55	48.45	21.55	27.02	26.90	24.52
Row players, last 30 rounds	49.05	50.95	25.95	25.48	25.00	23.57
Column players, last 30 rounds	54.05	45.95	17.14	28.57	28.81	25.48

null hypothesis of no difference in proportions is not rejected at the 5% level. It should be noted however, that because the sample size is very small the test lacks power and therefore requires the collection of more data for a definitive conclusion.

The difficulty in differentiating between the two types of heterogeneity is exacerbated by the fact that in human versus human experiments the experimenter in general does not have control over how one's opponent is playing and therefore a very large number of subjects would be required to have enough power to test between the two hypotheses. For a thorough discussion of these issues the reader is referred to Spiliopoulos (2008), which presents an experiment specifically designed to distinguish between within- and between-subjects heterogeneity.

5.7 Action profiles and convergence to the MSNE

The issue of convergence to MSNE was not at the core of N&S, although they do devote a small section to a pooled analysis where they discover that there appears to be a rough convergence toward the MSNE prediction of red play 60% of the time. Of particular interest is the question of whether players' choices were i.i.d. or whether there were two-period patterns that could possibly be exploited by a *gfp2* algorithm. Table 8 gives the percentages of play of single- and two-period sequences of actions both pooled for all players and broken down into column and row player subsets. The above variables are calculated both for all periods and also only for the last 30 periods in order to allow for learning and the possibility of convergence in the latter periods.

To test whether the proportions of first- and second-order play conform to the MSNE predictions χ^2 tests were performed using the actual frequencies of observed actions and the expected frequencies assuming a MSNE action profile. Pooled tests of all the players revealed that as a population it is possible to reject both first- and second-order MSNE behavior. Table 9 shows the individual results from the abovementioned χ^2 tests conducted at the 5% level (as well as belief model classifications which were addressed in Section 5.6). Each row in this table matches two opponents so that comparisons can be made across matched pairs of opponents e.g. player 1's opponent was player 7 in the experiment. Nine players were found not to be playing according to MSNE(1) - the probability of 9 out of 28 players players rejecting the null hypothesis at the 5% level given that the null hypothesis is true for all players is 0.000006. Twelve subjects were found to be deviating from MSNE(2) play and again the probability of obtaining at least this many rejections of the null hypothesis is close to zero. There are only three subjects who are found to deviate from second-order MSNE probabilities but at the same time do not deviate from first-order MSNE probabilities. Hence, it seems that players are not particularly prone to producing serially correlated behavior in this game despite the fact that the psychology literature has documented that subjects tend to produce serially correlated sequences when trying to generate a random sequence. However, the absence of serial correlation may not be because of explicit randomization on behalf of subjects but due to the interaction of deterministic learning rules, a possibility that is discussed in more detail in Section 5.9.

If these tests are restricted to the last thirty periods to allow for learning of first- and second-order MSNE probabilities in the first 30 periods, the results do not change much implying that many subjects have not been learning to play the MSNE. Seven of the players were found to be significantly deviating from MSNE(1) i.e. playing red 60% of the time and thirteen players were found to be deviating from i.i.d. behavior or MSNE(2) probabilities.

In both the 60 period and 30 period pooled analyses, the combination red-red is not played often enough whereas all other combinations are played more often than MSNE behavior would dictate. However an interesting picture emerges when the pooled analysis is broken down into two subsets: column and row players. In the last thirty periods of the game, column players chose red 54% of the time whilst row players only 49%. These differences then lead to even larger differences in second-order behavior of players. There is a marked difference regarding the play of green-green, as for column players it is 16.7% whereas for row players it is 26.2%. Interestingly, red-red play is extremely far away from the MSNE prediction of 36%, 25% for column players and 23% for row players[25].

Table 10 classifies each pair of players (along the rows of the table) based on the action profiles of both players. In five out of the fourteen pairs, MSNE action profiles are observed for both players and only one case exists where both players are only playing first-order MSNE probabilities but not second-order MSNE probabilities. There are six pairs where one player has a MSNE action profile but the other player deviates from it. In this case, the player who is deviating does not have an incentive to converge to a MSNE action profile because he attains MSNE payoffs no matter how he plays. However, the other player would have an incentive to change strategies, assuming that his opponent will not react to this. A possible explanation is that not all the players have the ability to detect these deviations by using a pattern recognition belief model and therefore are content playing a MSNE action profile which is a perfectly acceptable defensive strategy as they are not leaving themselves open to exploitation by a player who possibly uses pattern recognition. The problem of small sample size is again important in these comparisons - the Freeman and Halton (1951) extension of the exact Fisher test to 3×3 contingency tables does not reject the null hypothesis that there is no difference in proportions (p-value=0.41).

[25] However, these second-order probabilities are close to i.i.d. behavior given the observed first-order probabilities of play.

Table 9 Player classification in terms of best fitting belief model and action profiles

	Action profiles				Action profiles		
Player	MSNE(1)	MSNE(2)	Classification	Player	MSNE(1)	MSNE(2)	Classification
1	no	no	*gfp1*	7	yes	yes	*gfp2*
2	no	no	*gfp2*	8	yes	yes	*gfp1*/*gfp2*
3	no	no	*gfp2*	9	yes	yes	*gfp2*
4	no	no	*gfp1*	10	no	no	*gfp1*/*gfp2*
5	yes	no	*gfp2*	11	yes	no	*gfp2*
6	no	no	*gfp1*/*gfp2*	12	no	no	*gfp1*
13	no	no	*gfp1*	15	yes	yes	*gfp1*
14	yes	yes	*non-gfpn*	16	yes	no	*non-gfpn*
24	yes	yes	*gfp1*/*gfp2*	18	yes	yes	*gfp2*
25	yes	yes	*gfp2*	19	yes	yes	*gfp1*
17	no	no	*gfp1*	23	yes	yes	*gfp2*
20	yes	yes	*gfp1*	26	yes	yes	*gfp1*
21	yes	yes	*non-gfpn*	27	yes	yes	*gfp2*
22	yes	yes	*gfp1*	28	yes	yes	*gfp1*

Table 10 Classification of pairs based on action profiles

	∼MSNE(1)/∼MSNE(2)	MSNE(1)/∼MSNE(2)	MSNE(1)/MSNE(2)
∼MSNE(1)/∼MSNE(2)	2	0	5
MSNE(1)/∼MSNE(2)	0	1	1
MSNE(1)/MSNE(2)	5	1	5

Table 11 Payoffs of subjects grouped by classification

Row players				Column players			
non-gfpn	*gfp1*	*gfp2*	*gfp1*/*gfp2*	*non-gfpn*	*gfp1*	*gfp2*	*gfp1*/*gfp2*
4.225	4.2429	4.2292	4.45	3.6833	3.675	3.7917	3.7833

5.8 Simulated play of agents using different learning rules

The evolutionary fitness of the postulated learning algorithms can be examined by referring to the payoffs that each learning model can achieve when matched up against all the other learning models. In Table 11, the average payoffs of subjects are grouped by the learning model classification of each subject and by their role in the game as row or column players. However, due to the small sample size of the N&S experiment there do not exist enough observations in each group for statistical tests to have enough power to detect payoff differences. Also, any analysis of experimental data is necessarily restricted only to the learning models, and the associated parameter values, represented in the subject pool, so that some comparisons of interest might not be feasible to perform.

In response to these inadequacies, simulations were conducted where the two agents where programmed to play according to either *fp1* or *fp2* with a memory parameter of one, henceforth denoted as *fp1(1)* and *fp2(1)* (the number in the brackets denotes the value of the memory parameter) or *fp1* with a memory parameter of zero, *fp1(0)*, for 100 rounds in each game. As discussed earlier, in 2×2 games *fp1(0)* is equivalent to the win-stay/lose-shift heuristic (*ws*/*ls*). This is because *fp1(0)* assumes that an opponent's action in the current period will be the same as the previous period

Table 12 Statistics from simulations of *fp1(1)* versus *fp2(1)*

	fp1(1)		*fp2(1)*	
Players' mean statistics	Column	Row	Column	Row
Payoffs	3.492	4.12	3.88	4.5
$p(r)$	0.613	0.6	0.441	0.597
$p(g)$	0.387	0.4	0.559	0.403
$p(g-g)$	0.215	0.236	0.382	0.223
$p(g-r)$	0.168	0.161	0.162	0.166
$p(r-g)$	0.163	0.157	0.167	0.172
$p(r-r)$	0.434	0.43	0.27	0.42

action and then best responds to this. However, in games with a larger strategy space there will exist a problem because the lose-shift component of the heuristic does not prescribe which of the alternative actions the player should shift to.

The following variables were tracked during 1,000 simulations of each game: payoffs to each agent, first- and second-order probabilities of play[26]. The MSNE payoff for row players is 4.2 and for column players it is 3.8, with both row and column players expected to play red with probability 0.6.

5.8.1 Simulation of fp1(1) versus fp2(1) In simulations of these two agents, the *fp2(1)* agent had average payoffs higher than the MSNE payoffs (both when the *fp2(1)* agent was a column player and a row player), thereby necessarily imposing lower than MSNE payoffs upon the fp1(1) opponent, as can be seen in Table 12. In both cases the *fp1(1)* player exhibits strong serial autocorrelation as identified by $p(g-g)$ and $p(r-r)$ which are both greater than the MSNE prediction of 0.16 and 0.36 respectively. These deviations can then be detected and exploited by the *fp2(1)* player thereby explaining why the *fp2(1)* player can attain superior payoffs compared to the MSNE prediction at the expense of the *fp1(1)* player.

5.8.2 Simulation of fp1(0) versus fp2(1) The results differ significantly when the *fp1* player has a memory parameter of zero instead of one, as shown in Table 13. The *fp1(0)* player now manages to attain better than MSNE payoffs both as a column player as well as a row player to the detriment of the *fp2(1)* player. Both players' first- and second-order probabilities deviate from the MSNE prescription and they also end up playing the combination $g-g$ more often than the MSNE prescribes whilst playing all other two period combinations less often than the MSNE prescription.

5.8.3 Simulation of fp1(1) versus fp1(0) Table 14 shows that an *fp1(0)* agent does significantly better than the MSNE payoffs, both when playing as row and as column player. When the *fp1(1)* agent is a row player green is played twice in a row with probability 0.414 which is much higher than the MSNE prediction of 0.16. Hence, whenever the *fp1(0)* agent plays green and wins he will play green again which will now have a high probability of being his best response. When the perfect memory agent is a column player both red and green are repeated more often than they should be thereby again allowing the *fp1(0)* to have a higher success rate at playing his best response. The case where the row player is *fp1(0)* and the column player is *fp1(1)* is particularly interesting as the

[26] Small amounts of error were injected into a best response decision rule so as to generate some variability in actions and to avoid becoming mired in a single deterministic action profile.

Table 13 Statistics from simulations of *fp1(0)* versus *fp2(1)*

	fp1(0)		*fp2(1)*	
Players' mean statistics	Column	Row	Column	Row
Payoffs	3.9	4.333	3.667	4.1
$p(r)$	0.563	0.563	0.564	0.436
$p(g)$	0.437	0.437	0.436	0.564
$p(g-g)$	0.211	0.211	0.211	0.337
$p(g-r)$	0.216	0.216	0.215	0.216
$p(r-g)$	0.217	0.216	0.216	0.216
$p(r-r)$	0.336	0.337	0.338	0.211

Table 14 Statistics from simulations of *fp1(1)* versus *fp1(0)*

	fp1(0)		*fp1(1)*	
Players' mean statistics	Column	Row	Column	Row
Payoffs	4.118	4.517	3.483	3.882
$p(r)$	0.597	0.597	0.598	0.401
$p(g)$	0.403	0.403	0.402	0.599
$p(g-g)$	0.222	0.221	0.221	0.414
$p(g-r)$	0.172	0.173	0.172	0.172
$p(r-g)$	0.173	0.174	0.173	0.171
$p(r-r)$	0.413	0.412	0.414	0.222

first-order probabilities are equal to the MSNE(1) prediction of playing the red action 60% of the time. However, second-order play deviates from MSNE(2) predictions as instead of the probability of playing red twice in a row being 0.36 it is 0.221, which then leads to higher than MSNE payoffs for the row player. This highlights the necessity of examining second- and higher-orders of play because of the possibility of being misled into inferring MSNE play if only first-order probabilities of play are examined.

5.8.4 General observations from the agent simulations The first interesting result is that whether agents play first-order probabilities less than or greater than the MSNE probabilities may depend on whether an algorithm is playing as a row or column player. The second observation is that in the N&S game in most cases the payoff incentives to adopting the best possible learning model (out of the ones considered here) are not that significant as payoffs do not really increase much i.e. the curvature of the payoff function is relatively flat around the MSNE proportions. The changes in payoffs are larger when the row player is *fp2(1)* versus *fp1(1)* and in the two possible cases where an *fp1(0)* agent is playing an *fp1(1)* agent. From the strategies studied above, an *fp1(0)* agent, equivalent to the *ws/ls* heuristic, outperforms both of the other postulated models, *fp1(1)* and *fp2(1)*, and therefore, from the set of strategies examined, constitutes a best response strategy to these other models.

The good performance of a simple heuristic such as *ws/ls* may appear surprising however there exists well documented evidence from the psychology literature that such heuristics can in fact perform well compared to other complex rational models, in some cases even outperforming them. Martignon and Laskey (1999) argue that simple heuristics may perform better than complex models because the latter are vulnerable to overfitting on account of the large number of parameters that they have, a problem that is especially acute in very noisy environments. Also, the reduced number of free parameters of simple heuristics makes them more robust to variations in the environment. Another reason that heuristics may be effective is that they are tailored by evolutionary pressure to exploit structures in the environment. For example, the *ws/ls* heuristic is a very simple way of exploiting positively correlated events in the environment. For a further discussion of simple heuristics and their effectiveness/robustness the reader is referred to Gigerenzer (2000) and Gigerenzer and Selten (2001).

5.9 Are MSNE action profiles really the result of individual MSNE behavior?

An important distinction must be made between MSNE compatible action profiles and actual MSNE behavior on behalf of players. MSNE behavior by a player necessarily implies an expected MSNE action profile for that player, but the opposite is not necessarily true. MSNE compatible action profiles may be the result of deterministic learning processes by both players - it may simply be that the belief formation rule of one player when combined with the belief formation process of his opponent leads to apparent MSNE behavior. For example, pitting two *fp1(1)* agents or two *fp2(1)* agents against each other will lead to MSNE action profiles without either of the players actually behaving as a MSNE player. Also, as shown in Section 5.8.3 it is possible to observe MSNE(1) behavior from the interaction of two different deterministic learning models.

Whether MSNE action profiles are the result of MSNE behavior can be established by referring to Table 9. There are five cases where both players in a pair exhibit MSNE action profiles and in none of these cases is there a pair where both subjects are classified as *non-gfpn* players. In fact out of the ten subjects in these pairs only one player is classified as *non-gfpn*. Hence, it seems that MSNE action profiles are not the result of a conscious attempt to randomize but the result of the interaction of two *gfpn* belief formation or learning rules.

Concluding, there exist cases where even though players' beliefs are conditioned on the history of opponent's play (and therefore are not MSNE players) their resulting first-order action profile is consistent with MSNE. One should not be quick to assume that players are actually consciously trying to randomize in line with MSNE prescriptions just because their action profile appears to be consistent with MSNE behavior, as this may simply be the result of the two-way interaction of non-MSNE strategies of the players. Therefore the only way of definitely knowing whether this is the case or not is to directly examine the learning rules and decision processes employed by subjects.

5.10 Are stated beliefs better predictors of opponents actions than gfpn rules?

So far the discussion has focused on the value of *gfpn* learning rules in modeling elicited beliefs with the ultimate goal of these learning rules explaining the behavior of players. The focus now shifts from examining how well *gfpn* rules do in predicting own behavior to how well these rules do in predicting opponents' behavior. Table 15 documents how accurate various types of beliefs are in predicting opponents' behavior. The standard weighted fictitious play columns refer to the accuracy of standard *wfp* rules for prespecified, not estimated memory parameters. The estimated *gfpn* model columns refer to the beliefs that have been estimated on an individual basis and the stated beliefs row simply uses the beliefs elicited by N&S. In all cases, beliefs are assumed to predict that the opponent will play red if the probability belief is greater than 0.5, green if the belief probability

Table 15 Percentage of correct predictions of opponents' behavior

Standard weighted fictitious play						Estimated			
fp1			*fp2*			*gfpn* models			Stated
γ=0.5	γ=0.75	γ=1	γ=0.5	γ=0.75	γ=1	*gfp1*	*gfp2*	*ws/ls*	beliefs
51.35%	48.46%	57.64%	53.33%	55.48%	62.32%	53.45%	55.11%	55%	56.93%

is less than 0.5, and in cases where the belief is exactly equal to 0.5 it is assumed that the beliefs predict indifference.

Three interesting results were discovered. Firstly, out of the three standard models with fixed, non-estimated memory parameters, the best models were the ones with perfect memory. The speed of adjustment of weighted fictitious play beliefs depends on the length of history since they are essentially weighted averages. Hence, if only a few periods have passed the next observation will have a relatively large impact on the weighted average whereas after a large number of periods the impact will be smaller. The fact that the perfect memory model had the best performance can be interpreted in two ways. Either players maintained a constant strategy throughout the 60 rounds or if any significant changes in players' strategies do exist then they must occur early on when the fictitious play algorithm is able to adapt relatively quickly. Secondly, the standard weighted fictitious play models with no memory loss outperform stated beliefs in predicting opponents' behavior. The *fp2* model as would be expected outperforms the *fp1* model as it utilizes information about non-random behavior at a second-order probability level as well. Thirdly, for a given value of γ, the *fp2* learning rule always outperforms the *fp1* learning rule. This is further evidence that there exist exploitable patterns in the subjects' behavior and since *fp2* outperforms the stated beliefs it appears that subjects have not fully exploited these patterns.

The superiority of the standard *fp2* model with no memory decay is significant as it predicts opponent's behavior correctly 5.39% of the time more than stated beliefs do. One may then be tempted to ask why players do not adjust their belief formation process and use an *fp2*(1) model? It is important to remember that because of the bilateral interaction of the two players in this game if one player changes his belief formation model then this will result in a change in his actions which may then lead to a change in his opponent's actions. Hence, the answer is not as simple as looking at the success of belief formation models in predicting opponent's actions, since this implicitly assumes that opponents' actions are non-responsive to changes in beliefs of the other player. The answer can be explored by the simulations run in the previous section as these explicitly model the interaction between the two players and do not assume stationarity of one player's actions. Since a significant number of players use a *ws/ls* heuristic or equivalently *fp1(0)*, and from the simulations *fp2(1)* agents were found to earn less than MSNE payoffs when faced with *fp1(0)* agents, then the *fp2(1)* belief formation model is not necessarily the superior strategy to adopt in this particular population of subjects. This is a possible solution to the apparent paradox of why agents were not using an *fp2(1)* belief formation model instead of their stated beliefs.

6 Conclusion

This paper reexamined Nyarko and Schotter's seminal experimental research paper in which they directly elicited players beliefs instead of proxying them via some belief formation model. Modeling elicited beliefs with a standard weighted fictitious play model was not particularly successful, because the elicited beliefs time series was extremely volatile in contrast to the smoothness of weighted fictitious play beliefs. This paper proposed instead that players are capable of pattern recognition, which for the purposes of estimation was modeled by an appropriately modified generalized weighted fictitious play model, thereby allowing for a better fit to elicited beliefs. At the same time the parametric form of the learning rule was relaxed to include a constant and a coefficient of sensitivity to the fictitious play variables.

Results were encouraging as the fit of these models was significantly better than the fit of standard weighted fictitious play. Evidence was found that many of the players did use a pattern recognition belief formation rule. Specifically, ten players were classified as using two-period pattern detection models, or *gfp2* models, eleven players were found to use a weighted fictitious play algorithm, *gfp1* (the generalized analog of standard weighted fictitious play which includes a constant), only three were classified as exhibiting constant average beliefs with random fluctuations (*non-gfpn*) whilst four players were classified as *gfp1*/*gfp2* as there were insignificant differences in fit between these two models.

Also, it was found that Nyarko and Schotter's weighted fictitious play estimates of the memory parameter, γ, which they found to be centered on one, were probably biased because of the lack of a constant in these models. In actual fact, five out of eleven players classified as *gfp1* exhibited a memory parameter equal to zero, so that their behavior is equivalent to a generalized Cournot adjustment process, and the average memory parameter for all players classified as *gfp1* was 0.456. For players best modeled by *gfp2* algorithms, which employed pattern recognition, the average memory decay parameter was 0.678.

The inclusion of the constant and sensitivity parameters were strongly supported by the estimated models, as the average values were significantly greater than zero and significantly less than one respectively. This confirmed the hypothesis that players imperfectly incorporated standard weighted fictitious play variables in their beliefs as they exhibited less than perfect sensitivity to changes in these variables, probably due to bounded rational constraints on information processing and storage.

The finding of player heterogeneity as regards the incorporation of pattern detection into players' beliefs leads to the question of whether all subjects were capable of pattern detection but did not always employ it. This could occur because their opponent was randomizing efficiently and therefore there were no patterns to detect, or because some subjects have limited cognitive abilities and are incapable of pattern detection. The first explanation was coined as within-subjects heterogeneity, as it implies that agents have the ability to employ different models of behavior and choose which model to apply depending on the circumstances. The second explanation is between-subjects heterogeneity, which means that players do not have the same models of behavior at their disposal, perhaps due to different levels of bounded rationality. Some evidence was presented that players were more likely to employ two-period pattern detection in their beliefs when their opponents exhibited statistically significant deviations from i.i.d. behavior, thereby supporting the case that part of the discovered heterogeneity is within-subjects. This is an interesting result that warrants further investigation as the literature has not addressed the issue of the source of the heterogeneity that has been documented in experiments.

An investigation into the accuracy of various belief models in predicting an opponent's action led to the conclusion that a two-period pattern detecting model with perfect memory outperformed not only other models such as standard weighted fictitious play, but also performed better than subjects' stated beliefs. Hence, subjects were not optimally detecting patterns in opponents, and this is probably a result of the fact that players exhibited less than perfect recall or memory as measured by $\hat{\gamma}$ due to the increased complexity, memory and information processing necessary for pattern detection.

Although statistical tests of first- and second-order MSNE behavior concluded that many players were indeed adhering to the MSNE prescriptions, the belief formation models in nearly all cases rejected the hypothesis that their beliefs were MSNE beliefs. This is an important result as it shows that either this was the result of the interaction on the belief formation models of the players or that statistical tests on action profiles are not powerful enough to reject such a hypothesis, hence caution must be exercised in concluding that players are following the MSNE.

Finally, in cases where the sample size of the experimental data was not large enough to allow for powerful statistical tests, this paper reverted to simulations to examine behavior such as the evolutionary fitness of various belief learning models when pitted against each other. Employing different concurrent approaches is an effective method of bypassing the inherent limitations of the individual experimental techniques available to a researcher. Surprisingly, it was found that the simple win-stay/lose-shift heuristic outperformed fictitious play models both with and without pattern detection capabilities.

Further directions for research in this field would be to collect more experimental data with elicited beliefs to include games with larger action spaces and other types of strategic games with repeated interactions. These would increase the statistical power of tests of pattern detection thereby allowing the examination of tests of higher order pattern detection as well. Another possible direction is to follow a similar approach to the study by Shachat and Swarthout (2004), where subjects played against computer algorithms which were designed to deviate from the MSNE first-order prescriptions in an attempt to gauge how astute subjects were at detecting and exploiting these deviations. Modifying this experiment to include deviations in second- and higher order behavior would be extremely effective in unveiling the pattern detecting abilities of human subjects since the amount and types of deviations from i.i.d. behavior would be controlled by the experimenter.

A change in methodology to include neuroeconomic experiments would also be conducive to research. Data collected from such experiments could be crossmatched with neural research in the sequence learning literature to determine whether subjects are using the same areas of the brain to detect patterns in opponents. Other neuroeconomic studies (Platt and Glimcher, 1999; Dorris and Glimcher, 2004) have found that the activity of individual neurons in certain areas of the brain are strongly correlated with the value and likelihood of rewards, and even an approximate measure of expected value or utility, however these studies only varied first order-probabilities of reward. Again directly examining neuronal activity when second- and higher order probabilities are manipulated could provide direct evidence of the encoding of this type of information in the human brain.

References

Aizenstein, H., V. Stenger, J. Cochran, K. Clark, M. Johnson, R. Nebes, and C. Carter (2004). Regional Brain Activation during Concurrent Implicit and Explicit Sequence Learning. *Cerebral Cortex 14*, 199–208.

Bar-Hillel, M. and W. Wagenaar (1991). The perception of randomness. *Advances in Applied Mathematics 12*(4), 428–454.

Binmore, K., J. Swierzbinski, and C. Proulx (2001). Does Minimax Work? An Experimental Study. *Economic Journal 111*(473), 445–464.

Bloomfield, R. (1994). Learning a mixed strategy equilibrium in the laboratory. *Journal of Economic Behavior & Organization 25*(3), 411–436.

Blume, A., D. DeJong, G. Neumann, and N. Savin (1999). Learning in sender-receiver games. *University of Iowa working paper*.

Brown, J. N. and B. W. Rosenthal (1990). Testing the minimax hypothesis: A re-examination of o'neill's game experiment. *Econometrica 38*, 1065–81.

Cabrales, A. and W. Garcia-Fontes (2000). Estimating learning models from experimental data. *University of Pompeu Fabra working paper*.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C. F. and T. Ho (1999). Experience-weighted attraction learning in normal-form games. *Econometrica 67*, 827–74.

Cheung, Y. W. and D. Friedman (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior 19*, 46–76.

Chiappori, P. A., S. Levitt, and T. Groseclose (2002). Testing mixed strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review 92*(4), 1138–1151.

Cleeremans, A., A. Destrebecqz, and M. Boyer (1998). Implicit learning: news from the front. *Trends in Cognitive Sciences 2*(10), 406–416.

Clegg, B., G. DiGirolamo, and S. Keele (1998). Sequence learning. *Trends in Cognitive Sciences 2*(8), 275–281.

Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and their Application*. Cambridge University Press: Cambridge.

Destrebecqz, A., P. Peigneux, S. Laureys, C. Degueldre, G. Del Fiore, J. Aerts, A. Luxen, M. van der Linden, and A. Cleeremans (2003). Cerebral correlates of explicit sequence learning. *Cogn Brain Res 16*, 391–8.

Dorris, M. and P. Glimcher (2004). Activity in Posterior Parietal Cortex Is Correlated with the Relative Subjective Desirability of Action. *Neuron 44*(2), 365–378.

Fisher, R. (1920). A mathematical examination of the methods of determining the accuracy of observation by the mean error and the mean square error. *Monthly Notes of the Royal Astronomical Society 80*, 758–770.

Freeman, G. and J. Halton (1951). Note on an Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance. *Biometrika 38*(1/2), 141–149.

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

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. New York: Oxford University Press.

Gigerenzer, G. and R. Selten (Eds.) (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.

Gomez, R. L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology 33*, 154–207.

Gorard, S. (2005). The advantages of the mean deviation. *British Journal of Educational Studies 53*(4), 417–30.

Haruvy, E. and D. Stahl (2004). Deductive versus inductive equilibrium selection: experimental results. *Journal of Economic Behavior and Organization 53*(3), 319–331.

Huber, P. (1981). *Robust Statistics*. New York, John Wiley and Sons.

Kagel, John H. Roth, A. E. (Ed.) (1995). *The Handbook of Experimental Economics*. Princeton University Press.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*(2), 263–292.

Martignon, L. and K. Laskey (1999). *Simple Heuristics That Make Us Smart*, pp. 169–188. Oxford University Press.

Matheson, I. (1990). A critical comparison of least absolute deviation fitting(robust) and least squares fittings: the importance of error distributions. *Computers & chemistry 14*(1), 49–57.

Matlab (2007). *Mathworks, Inc., Natick, MA*.

McKelvey, Richard D. Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior 7*, 6–38.

Miller, G. (1956). The magical number seven, plus or minus two. *Psychological Review 63*, 81–97.

Mitchell, M. (1999). *An Introduction to Genetic Algorithms* (Fifth ed.). The MIT Press.

Mitropoulos, A. (2001). On the measurement of the predictive success of learning theories in repeated games. *Economics Working Paper Archive EconWPA*.

Nelder, J. A. and R. Mead (1965). A simplex method for function minimization. *Computer Journal 7*, 308–313.

Nissen, M. and P. Bullemer (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive psychology 19*(1), 1–32.

Nyarko, Y. and A. Schotter (2002). An experimental study of belief learning using elicited beliefs. *Econometrica 70*(3), 971.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. In *Proceedings of the National Academy of Sciences*, Volume 84, pp. 2106–9.

Palacios-Huerta, I. (2003). Professionals play minimax. *Review of Economic Studies 70*, 395–415.

Papke, L. E. and J. M. Wooldridge (1996, nov). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Platt, M. L. and P. Glimcher (1999). Neural correlates of decision variables in parietal cortex. *Nature 400*, 233–238.

Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics 117(3)*, 775–816.

Rapoport, A. and D. Budescu (1997). Randomization in individual choice behavior. *Psychological Review 104*(603-617).

Remillard, G. (2007). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *The Quarterly Journal of Experimental Psychology 1*.

Remillard, G. and J. M. Clark (2001). Implicit learning of First-, Second-, and Third-Order Transition Probabilities. *Journal of Experimental Psychology: Learning, Memory and Cognition 27*(2), 483–498.

Roth, A. E. and I. Erev (1995). Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior 8*(1), 164–212.

Salmon, T. C. (2001). An Evaluation of Econometric Models of Adaptive Learning. *Econometrica 69*(6), 1597–1628.

Shachat, J. and T. J. Swarthout (2004). Do we detect and exploit mixed strategy play by opponents? *Mathematical Methods of Operations Research 59*(3), 359–373.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association 62*, 626–633.

Sonnemans, J. and T. Offerman (2001). Is the Quadratic Scoring Rule Really Incentive Compatible? *University of Amsterdam manuscript*.

Spiliopoulos, L. (2008). Humans versus computer algorithms in repeated mixed strategy games.

Walker, M. and J. Wooders (2001). Minimax play at wimbledon. *American Economic Review*, 1521–38.

Wilson, H. (1978). Least squares versus minimum absolute deviations estimation in linear models. *Decision Sciences 9*, 322–335.

A Implications of the quadratic scoring rule

Assume one's opponent has two actions A or B to choose from and that $p(A)$ and $p(B)$ are the stated beliefs that the player reports. Whenever the opponent has played A the player will receive payoffs $0.1 - 0.05 \cdot [(1 - p(A))^2 + p(B)^2]$ however if B was chosen then the payoffs will be $0.1 - 0.05 \cdot [(1 - p(B))^2 + p(A)^2]$. Under this incentive scheme, expected payoffs are maximized by truthfully reporting one's beliefs. If players are risk neutral then this function is an optimal incentive mechanism because utility is maximized by stating beliefs truthfully. However, if players are risk averse then they might prefer to report beliefs near 0.5 thereby guaranteeing a minimum payoff. If on the other hand subjects were risk-seeking then it would be reasonable to expect near certain stated beliefs of zero or one. Sonnemans and Offerman (2001) offers a discussion of the quadratic scoring rule (QSR) and incentive compatibility in cases where risk preferences are not neutral or subjects are prone to probability weighting biases[27]. Although techniques exist for adjusting the QSR to retain its optimality property even in cases where risk preferences are not neutral, they require additional experimental procedures to obtain estimates with which to calibrate the adjusted QSR. Since such procedures were not undertaken in the original N&S paper it is impossible to apply such corrections in retrospect. Sonnemans and Offerman (2001) propose a method for correcting stated beliefs in light of non-risk neutral preferences, but fortunately conclude that these corrections did not adversely affect results in a number of studies they examined. They conclude that although it is desirable to make such corrections in future research, the influence is not strong enough to invalidate previous studies which did not use such corrections as the magnitude of the corrections are not large enough to significantly alter results.

B Computational details of model estimation and optimization

The problem of choosing initial parameter values for optimization algorithms is usually solved with one of two techniques.

1. Performing the optimization many times with randomly chosen parameter values from a restricted parameter space. The problem with this method is that for a large number of parameters it is necessary to use a large number of random initial points which means that an already computationally expensive optimization routine must be run a large number of times.
2. Performing a grid search over plausible parameter values and then choosing the best combination of parameter values (or a set of the best performing combinations) as initial starting points for the optimization algorithm. For a large number of parameters this technique is also computationally expensive as the number of possible initial grid values increases exponentially in the number of parameters (keeping the grid distance constant).

These two techniques are computationally inefficient because their search is not directed - in the first case points are chosen randomly, in the second case deterministically but arbitrarily. This occurs because the points are determined before the algorithm starts and are not updated with new information obtained during the execution of the algorithm. Incorporating this new information

[27] Probability weighting bias, first reported by Kahneman and Tversky (1979), refers to the empirical evidence that humans tend to overweight low probabilities and underweight high probabilities when making decisions.

would lead to an increase in efficiency as the algorithm could increase sampling from areas of the parameter space that show promise and are more likely to lead to better solutions. We propose instead the use of genetic algorithms where the search is influenced and directed by information collected throughout the procedure, see Mitchell (1999) for an introduction.

Genetic algorithms use three main evolutionary principles or operators to guide the search for an optimum.

1. Start with an initial (usually randomly selected) population of combinations of parameters and estimate the objective function for each one.
2. Apply a selection rule by selecting a subset of the best performing parameter combinations which will serve as the parents of the next generation.
3. Apply a crossover rule to combine two parents and create children for the next generation. This entails randomly selecting features from the two parents and combining them to form a new parameter combination.
4. Apply a mutation rule to chosen parents so as to randomly change features of the parent and create a child for the next generation.

Repeating the above three rules for each successive generation creates a process similar to natural or Darwinian selection. The selection rule guides the algorithm so that it spends more time in regions of parameter space which have been more successful in the past and therefore reduces the computational time devoted to scanning clearly suboptimal regions in the parameter space. The drawback is that it may wrongly become stuck in a suboptimal area. This is where the mutation rule is useful since it forces the algorithm to keep experimenting in other areas of the parameter space regardless of their past performance. The crossover rule allows for the process to quickly and efficiently hone in to a good solution by combining the characteristics of good performers.

Genetic algorithms are very well suited for problems that are discontinuous, nondifferentiable, stochastic, or highly nonlinear. Although quite often the genetic algorithm procedure is used by itself to perform the whole optimization routine, we propose to use it only to obtain initial parameter values. These will then be used to start another algorithm that is more suited to locally searching the parameter space and refining these initial parameter values found by the genetic algorithm. As argued in the main text, the family of gradient descent techniques may not be well suited to this specific problem and therefore the Nelder and Mead (1965) Simplex Method was implemented instead.