Available online at www.sciencedirect.com
SCIENCE DIRECT®

Games and Economic Behavior 51 (2005) 31–62

GAMES and Economic Behavior

www.elsevier.com/locate/geb

Learning, information, and sorting in market entry games: theory and evidence

John Duffy [a,*], Ed Hopkins [b]

[a] *Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA*
[b] *Edinburgh School of Economics, University of Edinburgh, Edinburgh EH8 9JY, UK*

Received 20 January 2003

Available online 18 August 2004

Abstract

Previous data from experiments on market entry games, N-player games where each player faces a choice between entering a market and staying out, appear inconsistent with either mixed or pure Nash equilibria. Here we show that, in this class of game, learning theory predicts sorting, that is, in the long run, agents play a pure strategy equilibrium with some agents permanently in the market, and some permanently out. We conduct experiments with a larger number of repetitions than in previous work in order to test this prediction. We find that when subjects are given minimal information, only after close to 100 periods do subjects begin to approach equilibrium. In contrast, with full information, subjects learn to play a pure strategy equilibrium relatively quickly. However, the information which permits rapid convergence, revelation of the individual play of all opponents, is not predicted to have any effect by existing models of learning.

JEL classification: C72; D83

Keywords: Market entry games; Reinforcement learning; Fictitious play; Experiments

* Corresponding author.
E-mail addresses: jduffy@pitt.edu (J. Duffy), e.hopkins@ed.ac.uk (E. Hopkins).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2004.04.007

1. Introduction

Theories of learning in games are increasingly being subjected to tests using data from controlled laboratory experiments with paid human subjects. The success or failure of various learning models has been assessed on the basis of how well these models predict or track the behavior of subjects in these experimental sessions. Given the usual short time horizon in experiments, researchers interested in testing models of learning have tended to concentrate on assessing their short-run fit. Long-run predictions have largely been ignored. One might reasonably be uncertain whether asymptotic results are likely to be relevant in experiments with finite length, or simply be interested in how subjects respond to novel situations. However, the long-run behavior of different learning models is often the same, giving clear hypotheses to test.[1]

This paper is a first attempt to see whether the *long-run* predictions of learning models do indeed help to explain behavior in the market entry game. This much studied game is a stylized representation of a very common economic problem: a number of agents have to choose independently whether or not to undertake some activity, such as enter a market, go to a bar, drive on a road, or surf the web, the utility from which is decreasing in the number of participants. Those choosing not to undertake the activity can be thought of as staying at home, staying out of the market, or simply not participating. Market entry games typically admit a large number of Nash equilibria. Pure equilibria involve considerable coordination on asymmetric outcomes where some agents enter and some stay out. The only symmetric outcome is mixed, requiring randomization over the entry decision. There also exist asymmetric mixed equilibria, where some agents play pure while others randomize. Given this multiplicity of equilibrium outcomes, an obvious question is: which type of equilibrium are agents likely to coordinate upon? Many previous experiments have been conducted in an effort to address this and related questions. See, for example, Rapoport et al. (1998, 2000, 2002), Seale and Rapoport (2000), Camerer and Lovallo (1999), Sundali et al. (1995), and Erev and Rapoport (1998). However, up to now, none of these studies has yielded evidence to suggest that repeated play leads to coordination on any type of Nash equilibrium, although in many experiments the average frequencies of entry in market entry games look remarkably like those generated by Nash equilibrium play.[2] That is, market entry games seem to represent a case where Nash equilibrium fails as a predictor of human behavior, at least at the individual level.

Here we investigate the alternative hypothesis that, given sufficient repeated play and adequate feedback, individuals in experimental market entry games should *learn equilibrium behavior*. This assertion leads naturally to further questions: what in practice is "sufficient" and what is "adequate"? How long should we expect to wait before agents coordinate on an equilibrium? What information is necessary? How do these factors interact, for example, does better information lead to faster convergence? In this paper, we attempt

[1] See, e.g. Hopkins (2002).

[2] But as Ochs (1998, p. 169) notes in a recent survey of experimental market entry game research, "... a common feature of all market game experiments ... is that the aggregate distributions [of entry rates] are not produced by Nash equilibrium profiles, that is, the *individual behavior* observed in all of these experiments is at variance with that implied by the best response conditions of a Nash equilibrium" (emphasis added).

to answer these questions in two ways. First, we provide formal results on long-run behavior in market entry games under two different models of learning that differ in terms of sophistication and use of information. Second, we report the results of a new series of experiments designed to test these predictions.

We show that two different models of learning predict not only that play should converge to a Nash equilibrium, but also that it should only converge to a subset of the total number of Nash equilibria. These predictions are in clear contrast with all previous experimental evidence on market entry games, which as noted above, has not been consistent with any Nash equilibrium. There are two models of learning which have attracted particular interest in explaining behavior in laboratory experiments, reinforcement learning and stochastic fictitious play.[3] They differ considerably in terms of sophistication. However, we show that under both, play must converge to an asymmetric pure equilibrium that involves what could be called "sorting," where some players always enter and the remaining players always stay out. However, these are asymptotic results. Thus, even if one of these learning models accurately describes human behavior, there is no guarantee that we would see the predicted outcome in the time available for laboratory experiments. What we seek to examine is whether such results are relevant in the timeframe of experiments, and by implication whether they are relevant outside the laboratory.

Previous experimental investigations of market entry games have concentrated on testing whether the symmetric mixed equilibrium or an asymmetric pure Nash equilibrium characterize the behavior of experimental subjects. In fact, the data seem to suggest a much more heterogeneous outcome, with some subjects apparently mixing between the two choices and some playing pure. However, the average number of entries per period is in rough accordance with equilibrium. Erev and Rapoport (1998) report two things of interest. First, distance from the symmetric mixed equilibrium is decreasing over time. Second, speed of convergence of the average number of entries toward Nash equilibrium levels is faster when more information is provided.

Learning models provide a potential explanation for the first of these experimental findings. For example, we show that under both reinforcement learning and stochastic fictitious play the mixed equilibrium is a saddlepoint, and hence movement toward this equilibrium in the short run is not inconsistent with convergence to a pure strategy equilibrium in the long run. In addition, Erev and Rapoport report a decrease in "alternation" over time, that is, the frequency that an agent plays the strategy which she did not play the previous period, which suggests individuals are getting closer to pure strategies. As to the second finding, the speed of convergence is more difficult to pin down theoretically and, in particular, the hypothesis that stochastic fictitious play that uses information about forgone payoffs is faster than simple reinforcement learning models that do not, has been difficult to formalize. Indeed, the results of our experiments are at variance with theoretical predictions about the impact of information on learning.

Existing experiments on market entry games have not provided ideal data sets to test the predictions of learning theory. For example, Rapoport et al. (1998) had sessions lasting

[3] There is now a very large literature. See for example Fudenberg and Levine (1998), Erev and Roth (1998), Camerer and Ho (1999).

100 periods, but within that time, the parameter which determined the number of entrants in equilibrium was constantly altered. Erev and Rapoport (1998) kept the parameters constant in each session, but each session lasted 20 periods, which is probably not sufficient for long-term behavior to emerge. As the capacity parameter c changes, the profile of strategies necessary to support a Nash equilibrium also changes, making coordination on a Nash equilibrium extremely challenging. There have been other experimental investigations of learning behavior employing large numbers of repetitions, for example, in Erev and Roth (1998), or in single person decision problems, Erev and Barron (2002). But the interest in these studies was to fit simulated learning models rather than to test theoretical results on convergence.

The new experiments on which we report here have several new features. First, each session involved 100 periods of play of an unchanging market entry game to give some chance for long-run behavior to be observed. Second, three different information treatments were employed. In the first "limited information" treatment, subjects were given no initial information about the game being played and each round were only told the payoff they earned. In the second, "aggregate information" treatment subjects were told the payoff function, and then were told after each round the number of subjects who had entered, the number who had stayed out, and the payoff each had earned. In the final "full information" treatment subjects were given the same information as in the aggregate information treatment, but in addition after each round the choice and payoff of each individual subject was revealed.

Our results are somewhat surprising. In the limited information treatment, there is some tendency for groups of subjects to converge upon a pure equilibrium, but only toward the very end of the 100 period session. The aggregate information treatment, despite the additional information provided, produced results very similar to those in the limited information treatment. In the full information treatment, the tendency toward sorting was much greater than in the other two treatments. This is despite the fact that all of the learning models considered predict no effect from the additional information provided in the full information treatment.

2. The game

The market entry game is a game with N players who must decide simultaneously and independently whether to enter a market or to stay out. One very simple formulation, found for example in Erev and Rapoport (1998), is where payoffs are linear in the number of entrants or participants. For example, if player i's strategy is $\delta^i = 0$ stay out, or $\delta^i = 1$ go in, then her payoff is

$$\pi_i(\delta) = \begin{cases} v, & \text{if } \delta^i = 0, \\ v + r(c - m), & \text{if } \delta^i = 1. \end{cases} \tag{1}$$

Here, v, r, c are positive constants and $0 \leqslant m \leqslant N$ is the number of agents that choose entry. The constant c, therefore, has the interpretation as the capacity of the market or road or bar. In particular, the return to entry exceeds the return to staying out, if and only if $m < c$. We can assume $1 \leqslant c < N$.

There are many pure strategy Nash equilibria for this class of games. If c is an integer, any profile of pure strategies which is consistent with either c or $c - 1$ entrants is a Nash equilibrium. If c is not an integer, a pure strategy Nash equilibrium involves exactly $\bar{c}$ entrants where $\bar{c}$ is the largest integer smaller than c. Moreover, if c is not an integer the number of Nash equilibria is finite, while if c is an integer there is a continuum of equilibria. The latter have the form, $c - 1$ players enter, $N - c$ stay out, and one player enters with any probability. Furthermore, this implies that only when c is not an integer are the pure equilibria strict.

Additionally, for $c > 1$, there is a symmetric mixed Nash equilibrium. This has the form

$$\bar{y}^i = \frac{c-1}{N-1} \quad \text{for } i = 1, \dots, N$$

where $\bar{y}^i$ is the probability of entry by the ith player. Note that the expected number of entrants in the symmetric mixed equilibrium is $c > N(c-1)/(N-1) > c-1$. There are additional asymmetric mixed equilibria[4] of the form $j < c - 1$ players enter with probability one, $k < N - c$ players stay out with probability one, and the remaining $N - j - k$ players enter with probability $(c - 1 - j)/(N - j - k - 1)$. In one of these asymmetric mixed Nash equilibria, the expected number of entrants is $j + (c - 1 - j)(N - j - k)/(N - j - k - 1)$ which again is between c and $c - 1$. Note though that as k approaches $N - c$, the expected number of entrants approaches c.

The common feature of all these Nash equilibria is that the expected number of entrants is between c and $c - 1$. This basic fact is corroborated by the experimental evidence. However, given the multiplicity of equilibria, it would be preferable if there were some way to select among the different equilibrium possibilities.

The market entry games that we examine here can be considered as one member of the large class of coordination games, characterized by having large numbers of Nash equilibria. However, unlike games of pure coordination, where players have an incentive for all to take the same action, here successful coordination involves different players taking different actions: some enter and some stay out. In one-shot play of such games, given that the players have identical incentives, one might think the symmetric equilibria is particularly salient, even though in this case it is mixed. In contrast, the insight from the literature on learning and evolution is that in repeated interaction, individuals will learn to condition their behavior on the behavior of others and hence converge to an asymmetric equilibrium. We go on to show that, in market entry games, under some well-known learning models, agents should indeed coordinate on a pure equilibrium, if only in the long run.

3. Models of learning

Here we identify two models of learning which differ in terms of the information they use but give the same clear prediction about how play in market entry games should develop in the long run. Imagine this game was played repeatedly by the same group of players at

[4] These asymmetric equilibria have not received much attention in previous experimental studies.

discrete time intervals $n = 1, 2, 3, \ldots$. We suppose that all players have propensities for the two possible actions. We write player i's propensities as (q^i_{1n}, q^i_{2n}), where here strategy 1 is entry and strategy 2 is staying out. Let the probability that agent i enters in period n be y^i_n and define $y_n = (y^1_n, \ldots, y^N_n)$. The probability of entry is determined by one of several possible mappings from propensities, for example, a linear choice rule

$$y^i_n = \frac{q^i_{1n}}{q^i_{1n} + q^i_{2n}}, \tag{2}$$

or the exponential rule

$$y^i_n = \frac{\exp(\beta q^i_{1n}/n)}{\exp(\beta q^i_{1n}/n) + \exp(\beta q^i_{2n}/n)}. \tag{3}$$

The principal focus of interest, however, is what information agents use in modifying their actions.

3.1. Simple reinforcement

Simple reinforcement is what is assumed in standard reinforcement learning models. That is, changes in propensities are a function only of payoffs actually received. In this case, the change in propensity for player i in period n would be

$$q^i_{1n+1} = q^i_{1n} + \delta^i_n\big(v + r(c - m_n)\big), \qquad q^i_{2n+1} = q^i_{2n} + \big(1 - \delta^i_n\big)v, \tag{4}$$

where m_n is the actual number of entrants in period n. This updating rule together with the choice rule (2) is what Erev and Roth (1998) call their basic reinforcement learning model. Note that given the choice rule (2), all propensities must remain strictly positive for the probability y^i to be defined. This can be assured, given the updating rule (4), if all payoffs are strictly positive. This last assumption is not usually problematic in an experimental setting, as experiments are usually designed, as were the ones reported here, so as not to give subjects negative payoffs.

3.2. Hypothetical reinforcement

In hypothetical reinforcement, in addition to undergoing simple reinforcement, an agent hypothesizes what she would have received if she had played strategies other than the one she actually chose. The payoff she would have received is then added to the corresponding propensity. In this context, this implies

$$q^i_{1n+1} = q^i_{1n} + v + r\big(c - m_n - \big(1 - \delta^i_n\big)\big), \qquad q^i_{2n+1} = q^i_{2n} + v. \tag{5}$$

Of course, use of this rule generally requires more information than simple reinforcement. Without knowledge of the payoff structure and the actions taken by opponents, it is very difficult to know what one would have received if one had acted differently. This updating rule together with the exponential choice rule (3) is an example of stochastic fictitious play (see for example, Fudenberg and Levine, 1998, Chapter 4).[5]

[5] Fictitious play is often modeled in terms of an agent having beliefs over the actions of opponents rather than in terms of propensities for his own actions. The two methods are largely equivalent (see, for example, Camerer

4. Learning dynamics

We now investigate the dynamics of the various types of learning introduced in the previous section. Each of these dynamics has differing requirements concerning information. Reinforcement learning requires only information on one own's payoff. Hypothetical reinforcement learning models such as stochastic fictitious play, require information about both the actions of others and the structure of payoffs. Thus there is an ordering of the two processes in terms of information requirements, which are reflected in the information treatments in our experiments. However, as we now show, the asymptotic behavior of these learning dynamics are not ordered in terms of their informational inputs. In fact, both reinforcement learning and fictitious play are predicted to lead to sorting.

To obtain some analytic results on the learning processes we consider, we make use of results from the theory of stochastic approximation. Simply put (see Appendix A for details), this allows investigation of the behavior of a stochastic learning model by evaluating its expected motion. In the case of the classic reinforcement learning process defined by the updating rule (4) and the choice rule (2), the expected motion of the ith player's strategy adjustment can be written as

$$E\left[y_{n+1}^i|y_n\right] - y_n^i = \frac{1}{Q_n^i} y_n^i\left(1 - y_n^i\right) r\left(c - 1 - \sum_{j\neq i} y_n^j\right) + O\left(\frac{1}{n^2}\right), \tag{6}$$

where $Q_n^i = \sum_j q_{nj}^i > 0$ is a player-specific scaling factor. Note that the right-hand side of the system of equations (6) is very close to the evolutionary replicator dynamics, which for this game would be the following system of differential equations:

$$\dot{y}^i = y^i\left(1 - y^i\right) r\left(c - 1 - \sum_{j\neq i} y^j\right) \tag{7}$$

for $i = 1, \ldots, N$.

The learning dynamics for market entry games when there are only two players are well understood. If capacity c is between 1 and 2, the game is similar to the games known as "Chicken" or "Hawk–Dove" in that, if you believe your opponent is going to adopt the "aggressive" strategy, enter, your best response is to stay out. There are three equilibria, a symmetric mixed equilibrium and two pure equilibria, where one of the two players enters and the other stays out. In this type of game, under adaptive learning any asymmetry is self-reinforcing as if one player initially enters with a high probability, then the other will move toward the best response of staying out entirely. Hence, under the replicator dynamics there will only be convergence to the mixed equilibrium if the initial conditions are entirely symmetric, that is, it is a saddlepoint, with the two asymmetric pure equilibria being stable attractors. Of course, when there are more than two players, there are many more equilibria including asymmetric equilibria where some agents randomize and some play pure and the dynamics are potentially much more complicated. However, we are able to show in Appendix A that in market entry games of arbitrary size, the behavior of the

and Ho, 1999; Hopkins, 2002) for two players but may differ in games with more than two players, depending on whether beliefs allow for correlation between the play of opponents.

replicator dynamics is similar to that in this simple case. We show, first, that the replicator dynamics (7) must converge to one of its rest points.[6] Second, we find that all mixed equilibria, symmetric or asymmetric, are saddlepoints and hence unstable.

This is a deterministic result. But this, together with stochastic approximation theory, allows us to show that reinforcement learning must also converge to a rest point of the replicator dynamics. The main difference is that, under the replicator dynamics (a deterministic process) if the system happened to start on the stable manifold of one the many saddlepoint mixed equilibria, then play would converge to that mixed equilibrium. However, if the actual stochastic learning process were to start on such a manifold, there is a zero probability of remaining in it, simply because of the noise implicit in the stochastic process. And if the learning process converges to an equilibrium but not to a mixed equilibrium, it must converge to a pure equilibrium. This is the intuition behind the following proposition, the proof of which is in Appendix A.

Proposition 1. *If agents use the reinforcement learning updating rule* (4) *and choice rule* (2)*, for generic values of c, with probability one, the learning process converges to a pure Nash equilibrium of the game. That is,* $\Pr\{\lim_{n\to\infty} y_n \in \overline{Y}\} = 1$, *where $\overline{Y}$ is the set of pure Nash equilibrium profiles.*

The reference to generic values of c refers to a difficulty mentioned earlier if c is an integer. In this case, there are an infinite number of Nash equilibria where $c-1$ agents enter with probability one, and $N-c$ agents stay out and with the remaining agent completely indifferent. Our intuition here about what a reasonable outcome constitutes and the analytic results available are both considerably weaker.

We now turn to hypothetical reinforcement and fictitious play. From Hopkins (2002), under the hypothetical updating rule (5) and the exponential choice rule (3), the expected motion of strategies can be written as

$$E\left[y_{n+1}^i | y_n\right] - y_n^i = \frac{\beta}{n+1}\left(y_n^i\left(1-y_n^i\right)r\left(c-1-\sum_{j\neq i} y_n^j\right) + \frac{1}{\beta}\sigma\left(y_n^i\right)\right) + O\left(\frac{1}{n^2}\right), \tag{8}$$

where $\sigma(y_n^i)$ is a noise term equal to

$$\sigma\left(y^i\right) = y_n^i\left(1-y_n^i\right)\left(\log\left(1-y_n^i\right) - \log y_n^i\right).$$

That is, the expected motion is close but not identical to the replicator dynamics. First, there is the additional noise term σ which ensures that each action will always be taken with a positive probability. Second, the expected motion is multiplied by the factor β. This has the effect that learning under stochastic fictitious play is much faster than under reinforcement learning.

[6] This is not as straightforward a result as it might seem. It is quite possible for dynamic systems, particularly in higher dimensions, to cycle and not converge to any single point.

The equilibrium points of such dynamics are not in general identical to Nash.[7] Define $\hat{y}$ as a perturbed equilibrium which satisfies for $i = 1, \dots, N$,

$$r\left(c - 1 - \sum_{j \neq i} y_n^j\right) + \frac{1}{\beta}\left(\log\left(1 - y^i\right) - \log y^i\right) = 0. \tag{9}$$

Note that for β sufficiently large and in generic games there will be a perturbed equilibrium for each Nash equilibrium. Second, as $\beta \to \infty$, the set of perturbed equilibria $\hat{y}$ approaches the set of Nash equilibria. Furthermore, by similar methods to those used in Proposition 1, we are able to establish the following result.[8]

Proposition 2. *If all players use hypothetical reinforcement together with the exponential choice rule* (3), *for generic values of c and for β sufficiently large, then* $\Pr\{\lim_{n\to\infty} y_n \in \widehat{Y}\} = 1$, *where $\widehat{Y}$ is the set of perturbed equilibrium each corresponding to one of the pure Nash equilibria of the game.*

5. Experimental design

The experimental design involves repeated play of the market entry game by a group of $N = 6$ inexperienced subjects under one of three different information conditions. We begin by discussing the parameterization of the payoff function and the three information conditions. We then explain the procedures followed.

We chose to set $v = 8$, $r = 2$ and $c = 2.1$ resulting in the following payoff function (in dollars):

$$\pi_i(\delta) = \begin{cases} \$8, & \text{if } \delta^i = X, \\ \$8 + 2(2.1 - m), & \text{if } \delta^i = Y. \end{cases}$$

where $0 \leqslant m \leqslant 6$ is the number of subjects (including i) choosing Y. We chose c to be non-integer so that, as noted, the number of Nash equilibria of the game would be finite and the pure equilibria would be strict. Nonetheless, we chose c to be close to an integer so that, similar to previous studies, in equilibrium there would be little difference in payoff to those entering and those staying out. The number of players, 6, is significantly smaller than in the previous experiments on market entry games. Our choice of $N = 6$ was based on the following considerations. We wanted a parameterization for the payoff function, in particular, a choice for the parameter r, that was similar to previous studies, and we wanted to provide subjects with reasonable compensation for their active participation. At the same time, we wanted to avoid any possibility that subjects earned *negative payoffs* that might result in ill-defined entry probabilities under the various learning models we examine.[9] These considerations favored our choice of a smaller number of subjects.

[7] See for example Fudenberg and Levine (1998, Chapter 4).

[8] A similar result for two player games is proved in Hofbauer and Hopkins (2002). Monderer and Shapley (1996) prove the convergence of fictitious play in this class of game.

[9] Erev and Rapoport (1998) use a parameterization of the payoff function that can result in significant negative payoffs given the larger number of subjects they consider (12). However, they adjust the propensity updating

In the first “limited information” treatment, subjects were repeatedly asked to choose between two actions X or Y, (which corresponded to “stay out” or “enter” respectively) without knowing the payoff function, π_i. Indeed, subjects did not even know that they were playing a game with other subjects. In this limited information treatment, each subject was informed only of the payoff from his own choice of action. Each subject’s history of action choices and payoffs was reported on their computer screens, and subjects also recorded this information on record sheets. Thus, in the limited information treatment, subjects had all the information necessary to play according to the reinforcement learning dynamic, but did not possess the information necessary for playing according to fictitious play.

In the second “aggregate information” treatment, subjects received feedback on the payoff from their action choice as in the limited information treatment, but were fully informed of the payoff function. In particular, subjects were told the payoff function, and to insure that their payoffs from choosing Y were as transparent as possible, the instructions also included the following table revealing all possible payoff values from choosing action Y. This table was also drawn on a chalkboard for all to see.

Fraction of players who choose action Y	1/6	2/6	3/6	4/6	5/6	6/6
Payoff each earns from choosing action Y	\$10.20	\$8.20	\$6.20	\$4.20	\$2.20	\$0.20

The instructions also clearly stated that the payoff each subject earned from choosing action X was always \$8, and this fact was also written on the chalkboard. Following the play of each round in the aggregate information treatment, subjects were further informed of the fraction of the six players who had chosen X and the fraction of the six players who had chosen Y, as well as the payoff received by all those choosing X and all those choosing Y. The past history (last 10 rounds) of the fractions choosing X and Y, along with the payoffs from each choice was always present on subjects’ computer screens, and subjects were asked to record this information on record sheets as well. Hence, in the aggregate information treatment, subjects had all the information necessary to play according to fictitious play.

In a final “full information” treatment, subjects were given all the information provided to subjects in the aggregate information treatment, and in addition, subjects were informed of the individual actions chosen by each of the other 5 players in the session, who were identified by their player ID numbers; this last piece of information was not available in the aggregate (or in the limited) information treatments. For example, as noted in the full information treatment instructions, the subject with ID number 3 might see that in the just completed round, the other 5 subjects’ choices were:

$$1X \;\; 2Y \;\; 4X \;\; 5X \;\; 6Y,$$

indicating that subject number 1 chose X, subject number 2 chose Y, subject numbers 4 and 5 both chose X and subject number 6 chose Y. The immediate past history (the last 10 rounds) of this individual action information was always present on subjects’ screens,

process of their learning model in the event that propensities become negative. It is less clear that the human subjects would make a similar adjustment.

thus enabling them to assess the extent to which the other 5 subjects were consistent or inconsistent in their choice of action. Since subjects in the full information all knew the payoffs earned each round by those choosing X and those choosing Y, they were provided a complete record of the actions chosen and payoffs earned by each individual subject in every round of the session.

We conducted nine one-hour sessions: three sessions for each of the three different information treatments. Each session involved exactly 6 subjects who had no prior experience playing the market entry game under any treatment (54 subjects total). Subjects were recruited from the undergraduate population at the University of Pittsburgh. In each session, the group of 6 subjects were seated at computer workstations, and were given written instructions which were also read aloud. Subjects were isolated from one another and no communication among subjects was allowed.

Subjects played the market entry game by entering their choice of action in each round, X or Y, using the computer keyboard when prompted by their monitor. Once all subjects had made their action choices, the computer program determined each subject's own payoff according to the parameterization of π_i given above, and reported this payoff back to each subject. Whether additional information was provided depended on the treatment as discussed above.

The six subjects played 100 rounds of the market entry game in an experimental session lasting one hour. Because the predictions that follow from Propositions 1–2 are all asymptotic, we wanted a sufficient number of repetitions to allow the predicted behavior to develop. Simulations of the various learning models (available on request) indicated that the 100 rounds allowed should be adequate at least for a pronounced movement toward equilibrium, if not actual convergence. Second, these simulations also indicated that, as learning slows over time, increasing the number of repetitions to 150, for example, would not produce radically different behavior.

The 100 rounds were broken up into four 25-round sets. Subjects were informed that at the end of each 25-round set, an integer from 1 to 25 would be randomly drawn from a uniform distribution with replacement. The chosen integer corresponded to one of the round numbers in the just completed 25-round set. Each subject's dollar payoff in that round was added to their total cash earnings for the session. This design was chosen to prevent subjects from becoming bored during the 100 repetitions of the market entry game. In addition to the 4 cash payments, subjects received \$5 for showing up on time and participating in the experiment. Average total earnings were \$37.87 in the limited information treatment, \$36.53 in the aggregate information treatment, and \$35.33 in the full information treatment.[10]

6. Equilibrium predictions and hypotheses

Given our parameterization of the market entry game, pure strategy Nash equilibria have 2 players always entering, each earning \$8.20, and 4 players always staying out, each earn-

[10] These amounts include the \$5 payment and the four randomly determined payoff amounts. Average per round payoffs are reported in Section 7.1.

Table 1
Equilibrium predictions

Equilibrium	Number of entrants	
	Mean	Standard deviation
Pure	2	0
Symmetric mixed	1.32	1.015
Asymmetric mixed	1.467	0.964
Pure QRE ($\beta = 5$)	1.781	0.512
Symmetric QRE ($\beta = 5$)	1.457	1.050
Asymmetric QRE ($\beta = 5$)	1.525	0.968

ing \$8.00. The unique symmetric mixed strategy Nash equilibrium prediction is that each player enters with probability 0.22 and earns an expected payoff of \$8.00. In this equilibrium, the expected number of entrants is 1.32. Finally, as noted in Section 3, there are many asymmetric mixed equilibria. However, play in some of the sessions seems to approach one of these in particular. In this asymmetric mixed equilibrium, 2 players always stay out and the remaining 4 players enter with probability 0.367, earning an expected payoff of \$8.00 each. The expected number of entrants in this asymmetric mixed equilibrium is 1.467. As noted, if subjects were to use a perturbed choice rule such as the exponential rule (3), the steady states of the learning process would not be Nash equilibria, but perturbed equilibria (also known as Quantal Response Equilibria (QRE) after McKelvey and Palfrey, 1995). We report also the QRE equilibria (for a typical value of the parameter β) that correspond to the three Nash equilibria of interest. These equilibrium predictions are summarized in Table 1.

Together with the theoretical results of the previous section, we can make the following hypotheses.

Hypothesis 1. If subjects are reinforcement learners, then:

(a) play should evolve over time toward a pure strategy Nash equilibrium, and
(b) there should be no change in the speed with which play evolves toward this equilibrium in the limited information treatment as compared with the aggregate information or full information treatments.

Hypothesis 2. If subjects are hypothetical reinforcement learners, playing according to stochastic fictitious play or a threshold learning rule, then

(a) play should evolve over time toward a (perturbed) pure strategy Nash equilibrium in the aggregate and full information treatments;
(b) there should be no change in the speed with which play evolves toward a pure strategy equilibrium in the aggregate information treatment as compared with the full information treatment.

Note that the fictitious play requires information that was not made available to subjects in our limited information treatment. It is therefore unclear what this model predicts in such circumstances. There has been more than one attempt (for example, Fudenberg and Levine, 1998, Chapter 4; Sarin and Vahid, 1999; Anderson and Camerer, 2000) to specify a fictitious play-like learning process for environments where opponents' play is not observable. However, the properties of these learning processes, and in particular, their speed with respect to fictitious play, are not well known. Therefore, we treat fictitious play as making no predictions in the limited information treatment.

It has been suggested to us that, given that expected payoffs are similar in both the pure and mixed equilibria, there are not sufficient incentives for subjects to learn to play a pure equilibrium. This similarity is inevitable given that in any equilibrium of this game, both actions must be active and expected payoffs be sufficiently close so as to give no incentive to deviate. This observation is reflected in the learning dynamics outlined in Section 4. As noted, although all mixed equilibria are unstable, they are saddlepoints, which implies there is at least one stable path leading to each equilibrium. If play starts near such a path, then the learning dynamics may take some considerable time to move away from the neighborhood of the equilibrium. Therefore, although learning theory predicts convergence to a pure equilibrium, such play may take a long time to emerge.

7. Experimental findings

7.1. Main results

The main findings are summarized in Tables 2–4 and Fig. 1. Table 2 reports the session–level means and standard deviations in per round payoffs over all 100 rounds, over the last 50 rounds and over the last 10 rounds of each session. Table 3 does the same for a related measure, the number of players choosing to enter. Figure 1 reports the round-by-round mean number of entrants across the three sessions of each treatment, along with a one standard deviation bound. Finally, Table 4 reports *individual* subject entry frequencies and standard deviations.

Table 2
Mean and standard deviation in per round payoff (in dollars) over all 100 rounds, the last 50 rounds and last 10 rounds of each session

	Session #1, rounds:			Session #2, rounds:			Session #3, rounds:		
	All 100	Last 50	Last 10	All 100	Last 50	Last 10	All 100	Last 50	Last 10
Limited info.									
Mean	7.85	7.90	8.13	7.84	7.94	8.03	7.76	7.90	7.78
St. dev.	0.95	0.96	0.80	0.98	0.82	0.58	1.03	0.74	0.68
Aggregate info.									
Mean	7.65	7.79	7.77	7.68	7.73	7.78	7.47	7.53	6.70
St. dev.	1.14	0.95	0.93	1.08	0.80	0.67	1.30	1.17	1.19
Full info.									
Mean	7.83	8.05	8.07	7.71	7.87	8.07	7.70	7.75	7.93
St. dev.	0.63	0.24	0.11	1.09	0.81	0.10	1.10	0.90	0.71

Table 3
Mean and standard deviation in the number of entrants over all 100 rounds, the last 50 rounds and last 10 rounds of each session

	Session #1, rounds:			Session #2, rounds:			Session #3, rounds:		
	All 100	Last 50	Last 10	All 100	Last 50	Last 10	All 100	Last 50	Last 10
Limited info.									
Mean	1.65	1.54	1.10	1.92	1.86	1.90	2.10	2.08	2.30
St. dev.	1.09	1.08	0.83	0.91	0.80	0.54	0.84	0.60	0.46
Aggregate info.									
Mean	2.13	2.04	2.10	2.17	2.32	2.30	2.39	2.40	3.00
St. dev.	0.99	0.87	0.83	0.90	0.55	0.46	0.94	0.83	1.10
Full info.									
Mean	2.21	2.00	2.00	2.18	2.06	2.00	2.20	2.22	2.00
St. dev.	0.52	0.20	0.00	0.83	0.68	0.00	0.82	0.67	0.63

Table 2 reveals that per round payoffs are similar across the three sessions of each treatment.[11] Closer inspection reveals that over all 100 rounds, the mean per round payoffs in the three aggregate information sessions are significantly lower (at the 5% level of significance) than the comparable means for either the three limited or the three full information sessions according to a nonparametric robust rank order test. However the difference in mean payoffs between the aggregate and full information sessions becomes insignificant once attention is restricted to the last 50 rounds of these sessions. There is no significant difference in the mean payoffs between the limited and full information treatments over any of the horizons reported in Table 2.

In Table 3 we see that a related aggregate statistic—the mean number of entrants over all 100 rounds—is lower in the limited information treatment as compared with either the aggregate or full information treatments. These differences are significant at the 5% level, again using the robust rank order test. However, these differences becomes insignificant once attention is restricted to the last 50 rounds of a session.[12] Table 3 as well as Fig. 1 reveals that in all three treatments, the mean number of entrants generally lies between c and $c-1$, or between 2.1 and 1.1, though there are some exceptions. In particular, over the last 50 rounds of two of the three aggregate information sessions and one of the three full information sessions, the average number of entrants exceeded 2.1 by small amounts. Perhaps the most interesting finding in Table 3 and Fig. 1 is that in each of the three full information treatments, there appears to be perfect coordination on a pure Nash equilibrium for at least one 10-round period, i.e. the standard deviation for that 10-round entry frequency was zero (see, in particular Fig. 1). Of course, to assess whether a pure Nash equilibrium was actually achieved requires further disaggregation of the data, which is done in Table 4.

[11] Recall that subjects were only paid on the basis of four randomly chosen rounds, so the payoff means reported in Table 2 (over 100 rounds, the last 50 rounds, and the last 10 rounds) are not the same as actual mean payoffs.
[12] We also note that, according to a Kolmogorov–Smirnoff test, there is no significant difference in the distribution of the initial (i.e. first round) number of entrants between any two of the three treatments (using the three initial distributions available for each treatment).

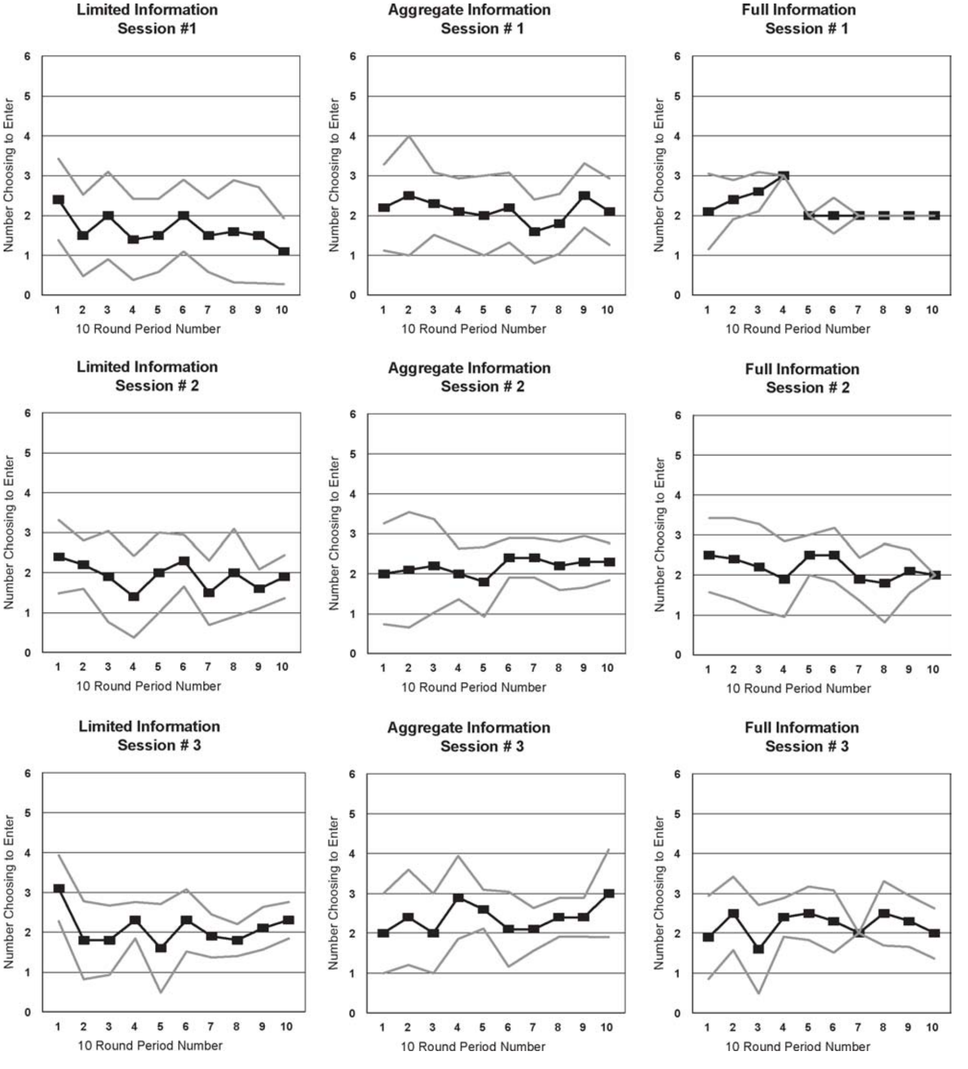


Fig. 1. 10-round mean number of entrants, one-standard deviation bound.

This table reports the mean and standard deviation of the entry frequencies for each subject in every session. Looking at the full information treatment results, we see that in two of the three sessions (full information sessions numbers 1 and 2), subjects did indeed achieve perfect coordination on the pure equilibrium where 2 players always enter and 4 always stay out over the last 10 rounds of these sessions, as the standard deviation of the entry frequencies are zero for each subject.

We note further that in full information session 1, subjects actually achieved a pure strategy Nash equilibrium much earlier, from rounds 41–51, and another pure strategy

Table 4
Individual entry frequencies: means and (standard deviations) over all 100 rounds, the last 50 rounds and last 10 rounds of each session

Player number	Session #1			Session #2			Session #3		
	All 100	Last 50	Last 10	All 100	Last 50	Last 10	All 100	Last 50	Last 10
Limited info.									
1	0.07	0.08	0.00	0.76	0.84	1.00	0.12	0.02	0.00
	(0.26)	(0.27)	(0.00)	(0.43)	(0.37)	(0.00)	(0.32)	(0.14)	(0.00)
2	0.26	0.36	0.10	0.58	0.64	0.80	0.04	0.00	0.00
	(0.44)	(0.48)	(0.30)	(0.49)	(0.48)	(0.40)	(0.20)	(0.00)	(0.00)
3	0.34	0.32	0.30	0.04	0.02	0.00	0.67	0.76	0.80
	(0.47)	(0.47)	(0.46)	(0.20)	(0.14)	(0.00)	(0.47)	(0.43)	(0.40)
4	0.56	0.58	0.60	0.00	0.00	0.00	0.24	0.30	0.50
	(0.50)	(0.49)	(0.49)	(0.00)	(0.00)	(0.00)	(0.43)	(0.46)	(0.50)
5	0.03	0.00	0.00	0.25	0.02	0.00	0.13	0.04	0.00
	(0.17)	(0.00)	(0.00)	(0.43)	(0.14)	(0.00)	(0.34)	(0.20)	(0.00)
6	0.39	0.20	0.10	0.29	0.34	0.10	0.90	0.96	1.00
	(0.49)	(0.40)	(0.30)	(0.45)	(0.47)	(0.30)	(0.30)	(0.20)	(0.00)
Aggregate info.									
1	0.01	0.00	0.00	0.22	0.14	0.10	0.76	0.76	0.60
	(0.10)	(0.00)	(0.00)	(0.42)	(0.35)	(0.30)	(0.43)	(0.43)	(0.49)
2	0.53	0.46	0.50	0.49	0.74	0.20	0.10	0.00	0.00
	(0.50)	(0.50)	(0.50)	(0.50)	(0.44)	(0.40)	(0.30)	(0.00)	(0.00)
3	0.60	0.54	0.70	0.15	0.00	0.00	0.04	0.04	0.00
	(0.49)	(0.50)	(0.46)	(0.36)	(0.00)	(0.00)	(0.20)	(0.20)	(0.00)
4	0.61	0.60	0.20	0.00	0.00	0.00	0.29	0.30	0.50
	(0.49)	(0.49)	(0.40)	(0.00)	(0.00)	(0.00)	(0.46)	(0.46)	(0.50)
5	0.00	0.00	0.00	0.36	0.44	1.00	0.70	0.76	0.90
	(0.00)	(0.00)	(0.00)	(0.48)	(0.50)	(0.00)	(0.46)	(0.43)	(0.30)
6	0.38	0.44	0.70	0.95	1.00	1.00	0.50	0.54	1.00
	(0.49)	(0.50)	(0.46)	(0.22)	(0.00)	(0.00)	(0.50)	(0.50)	(0.00)
Full info.									
1	0.35	0.02	0.00	0.02	0.00	0.00	0.73	0.96	1.00
	(0.48)	(0.14)	(0.00)	(0.14)	(0.00)	(0.00)	(0.44)	(0.20)	(0.00)
2	0.05	0.00	0.00	0.46	0.74	1.00	0.00	0.00	0.00
	(0.22)	(0.00)	(0.00)	(0.50)	(0.44)	(0.00)	(0.00)	(0.00)	(0.00)
3	0.81	1.00	1.00	0.25	0.08	0.00	0.00	0.00	0.00
	(0.39)	(0.00)	(0.00)	(0.43)	(0.27)	(0.00)	(0.00)	(0.00)	(0.00)
4	0.01	0.00	0.00	0.51	0.16	0.00	0.70	0.64	0.70
	(0.10)	(0.00)	(0.00)	(0.50)	(0.37)	(0.00)	(0.46)	(0.48)	(0.46)
5	0.01	0.00	0.00	0.67	0.96	1.00	0.51	0.52	0.30
	(0.10)	(0.00)	(0.00)	(0.47)	(0.20)	(0.00)	(0.50)	(0.50)	(0.46)
6	0.98	0.98	1.00	0.27	0.12	0.00	0.26	0.10	0.00
	(0.14)	(0.14)	(0.00)	(0.44)	(0.32)	(0.00)	(0.44)	(0.30)	(0.00)

equilibrium beginning in round 54; they remained in the latter pure strategy equilibrium for the last 46 rounds of the experiment (see Fig. 1). In full information session 2, subjects achieved a pure strategy equilibrium in round 85 and remained in that equilibrium for the last 15 rounds of the experiment. In full information session 3, a pure strategy equilibrium was achieved only briefly from rounds 63–69 (7 rounds).[13] However, we note that by the last 10 rounds of full information session 3 four of the six players were adhering to pure strategies; one always in and three always out.

Table 4 reveals that there is some support for Hypothesis 1(a): as the reinforcement learning model predicts, subjects in the limited information sessions are close to coordinating on a pure equilibrium by the end of each of the three limited information sessions. Note that by the final 10 rounds of each session, three or four players choose not to enter at least 90% of the time, and one or two players choose to enter more than 50% of the time. Moreover we see that the standard deviations for the individual entry frequencies are almost always lower in the last 10 rounds as compared with the last 50 rounds. On the other hand, there does not appear to be much support for Hypothesis 1(b) as there are differences in the speed of convergence as subjects are given more information in the aggregate information and full information treatments. In particular, convergence toward the pure strategy equilibrium appears to be much faster in the full information treatment as compared with the limited or aggregate information treatments.

In the last 20 rounds of the three aggregate information sessions, subjects appear to be somewhere between the asymmetric mixed equilibrium and the pure equilibrium. That neither equilibrium has been reached is supported by the fact that there is excessive entry relative to that predicted in either equilibrium (compare the mean number of entrants in Table 3 with the predictions in Table 1). Notice also in Table 4 that in the last 50 (and last 10 rounds) of each of the three aggregate information treatments there remain four players who are still choosing to enter with some positive frequency, and exactly two players who (almost) purely stay out.

While Table 4 is informative about individual behavior, the individual frequencies and standard deviations provide somewhat imprecise evidence regarding the closeness of play by each group of 6 subjects to the pure strategy prediction. To get further at this issue, we make use of the Gini index of inequality. Let P_i be the percentage of all decisions to enter ($\delta^i = 1$) made by player i over R rounds, (e.g. the last 50 rounds of a session): $P_i = N_i/N$, where $N_i = \sum_{j=1}^{R} \delta^i_j = 1)$ and $N = \sum_{i=1}^{6} N_i)$. If two players always enter and the remainder always stay out over the R-round interval, the vector of P_i values, sorted from least to most is $P = \{0.0,\ 0.0,\ 0.0,\ 0.0,\ 0.5,\ 0.5\}$, and the Gini coefficient is equal to 0.667.[14] By contrast, if all players were entering an equal percentage of the time

[13] Figure 1 may give the mistaken impression that a pure strategy Nash equilibrium was obtained over rounds 61–70 of full information session 3. In fact, there were just two entrants in each round of this 10-round interval, but in rounds 63 and 70, one subject who had been an entrant in previous rounds chose not to enter while another subject who had been staying out simultaneously chose to enter. Hence, the standard deviation in the number of entrants was indeed 0 over rounds 61–70, as reported in Fig. 1, but a pure equilibrium was only obtained over the shorter interval consisting of rounds 63–69.

[14] The Gini coefficient is defined as:

$$G = \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K}\bigl|(1/K)P_i - (1/K)P_j\bigr|,$$

where K is the number of components; in our case $K = 6$ individuals. Note that unlike the mean squared deviation criterion discussed below, in Section 7.2, the Gini coefficient does not require a determination of *which* players are playing certain pure or mixed strategies.

Table 5

Gini coefficient treatment–session	Rounds		
	All 100	Last 50	Last 10
Low info.-1	0.373	0.418	0.591
Low info.-2	0.474	0.572	0.685
Low info.-3	0.481	0.583	0.572
Mean (all low)	0.443	0.524	0.616
Agg. info.-1	0.389	0.379	0.468
Agg. info.-2	0.454	0.540	0.587
Agg. info.-3	0.391	0.431	0.433
Mean (all agg.)	0.411	0.450	0.496
Full info.-1	0.569	0.663	0.667
Full info.-2	0.323	0.552	0.667
Full info.-3	0.455	0.536	0.617
Mean (all full)	0.449	0.584	0.650

($P_i = 1/6$), as in the symmetric mixed strategy equilibrium, the Gini coefficient would be 0. As Table 5 reveals, the Gini coefficients come closest to the pure strategy predicted value of $2/3$ in the full information treatment and come least close to the $2/3$ prediction in the aggregate information treatment.

The difference in findings between the limited and full information treatments appear to lie in the speed of convergence and not the type of equilibrium selected. In particular, it seems that additional information may affect the speed of convergence to a pure strategy equilibrium in violation of the notion that subjects are strictly reinforcement learners. So reinforcement learning does reasonably well in explaining behavior in the low information treatment, in that even there is movement toward equilibrium. But given that it fails to pick up the effect from greater information, models that use more information might fit better in the aggregate and full information treatments.

There is much more support for Hypothesis 2(a) than for 2(b). Whereas play does seem to approach a pure strategy equilibrium in the aggregate and full information treatments, it also appears that the additional information provided in the full information treatment relative to the aggregate information treatment has a substantial effect on subject behavior; subjects in the full information treatment are much closer to the pure strategy equilibrium by the end of the session than are subjects in the aggregate information treatments; indeed, as noted earlier, in two of the three full information sessions subjects had achieved and sustained perfect coordination on a pure equilibrium by the end of the session. Finally, all of the sessions give greater support for the Nash equilibrium of the one-shot game than for collusive strategies aimed at maximizing joint payoffs. Indeed, such strategies, which involve all subjects taking turns to be the sole entrant, involve an implausible level of coordination, and have not been observed in previous experimental studies.

7.2. Convergence to equilibrium

To determine how close subjects were to convergence on a particular equilibrium, we first calculated each subject's entry frequency over 10-period, non-overlapping samples,

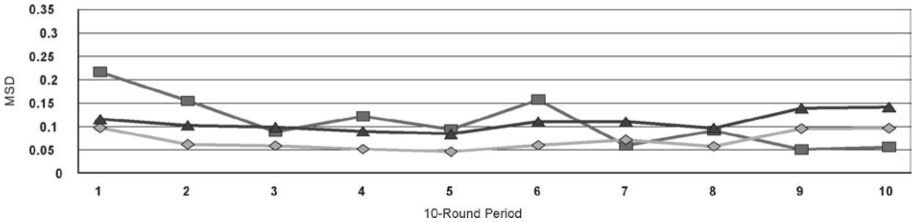
$s = 1, 2, \ldots, 10$. Denote the entry frequency of subject i over sample s by y_s^i. We then calculated the mean squared deviation (msd) from a predicted equilibrium entry frequency $\hat{y}^i$, over each 10-period sample in session j, $\text{msd}_s^j = 1/6 \sum_{i=1}^{6} (y_s^i - \hat{y}^i)^2$. This calculation is straight–forward for the unique symmetric mixed equilibrium, where $\hat{y}^i = 0.22$ for all i. Since there are many pure and asymmetric mixed equilibria, we chose to select *one* equilibrium of each type for each session. Each pure equilibrium was selected by determining the two players who were closest to playing the pure strategy of always entering over the last 10 rounds of each session. The other four players were regarded as being closest to playing the pure strategy of always staying out. This assignment of pure strategies was then used over all periods of the session (i.e. starting from period 1). In all sessions, the assignment of pure strategies to players based on final 10-round entry frequencies was readily apparent from Table 4.[15] Depending on this categorization, the predicted entry frequency, $\hat{y}^i$, would be either 1 or 0, and using these predictions, we calculated the msd from "the" pure strategy for each s in each session. Similarly for the asymmetric mixed equilibrium, we used the final 10-round and sometimes the final 50-round entry frequencies to determine the two players in each session who were closest to playing the pure strategy of always staying out, $\hat{y}^i = 0$. The other four players were regarded as being closest to playing the mixed strategy which has a predicted entry probability of $\hat{y}^i = 0.367$.[16] Again, these assignments were in most cases, readily apparent, and the assignment of players to pure or mixed strategies was used over all periods of the session.[17]

Figure 2 shows the sequence of 10-period, mean squared deviations averaged over the three sessions of each information treatment, $(1/3 \sum_{j=1}^{3} \text{msd}_s^j)$. In all three information treatments, the (average) msd from "the" pure equilibrium is initially much higher than the msd from the other two types of equilibria, but by the final 10 rounds, the msd from the pure equilibrium is less than the msd from these other equilibrium types. In the case of the full information treatment, the msd from the pure equilibrium falls below the msd from the other equilibrium types between periods 50–60, and remains there for the duration of the full information sessions. Notice also that in the aggregate and full information treatments, the msd from the asymmetric and the symmetric mixed equilibria appears to be rising over time.

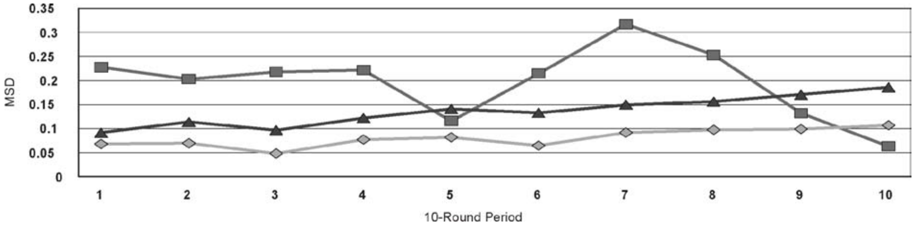
[15] Players deemed to be playing the pure strategy of always entering were (session: player numbers): Lim. info. #1: 3,4; Lim. info. #2: 1,2; Lim. info. #3: 3,6; Agg. info. #1: 3,6; Agg. info. #2: 5,6; Agg. info. #3: 5,6; Full info. #1: 3,6; Full info. #2: 2,5; Full info. #3: 1,4. The rest were deemed to be playing the pure strategy of staying out.

[16] Those deemed to be playing the pure strategy of staying out were (session: player numbers): Lim. info. #1: 1,5; Lim. info. #2: 3,4; Lim. info. #3: 1,2; Agg. info. #1: 1,5; Agg. info. #2: 3,4; Agg. info. #3: 2,3; Full info. #1: 4,5; Full info. #2: 1,3; Full info. #3: 2,3. The rest were deemed to be playing the mixed strategy.

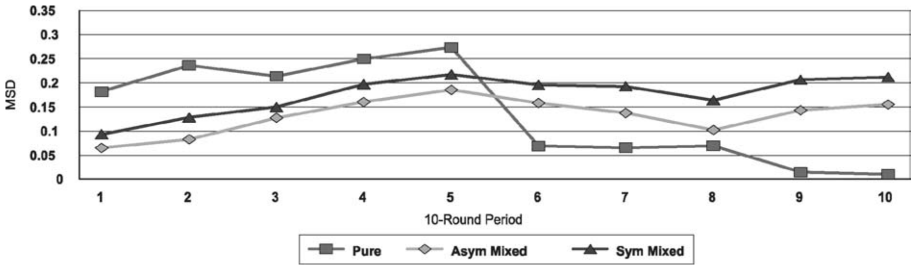
[17] We recognize that msd can be an imperfect measure of convergence to a mixed strategy equilibrium as it cannot detect sequential dependencies in players' entry choices. However, since we do not find that players converge to a mixed strategy or asymmetric mixed equilibrium using our msd convergence criterion, it seems unlikely that alternative convergence criteria that were capable of detecting sequential dependencies in entry choices would alter our findings.



(a) Limited information.



(b) Aggregate information.



(c) Full information.

Fig. 2. 10-round mean squared deviations from the three types of equilibria: averages over all 3 sessions of a treatment.

7.3. *Information and learning*

While it appears that the amount of information that subjects are given affects their behavior, we have yet to provide direct evidence that subjects are reacting differently to the different types of information they are given or whether subjects can be properly characterized as reinforcement learners or as hypothetical reinforcement learners in those treatments where hypothetical reinforcement is possible. We begin by considering whether subjects condition their entry decision on information concerning the number of entrants. Figure 3 shows the average frequency of entry by all six members of a group in period n conditional on the number of entrants in that same group in period $n-1$, using data averaged over all three sessions of a treatment. Attention is restricted to the case of 0–4 entrants as there were only two periods in all 9 sessions where more than 4 players entered (both occurred in the aggregate information treatment). The numbers on each bar indicate the fraction of observations for that treatment falling into that bin, e.g. 25% of the observations from the three limited information sessions were for the case where 1 player entered in period $n-1$.

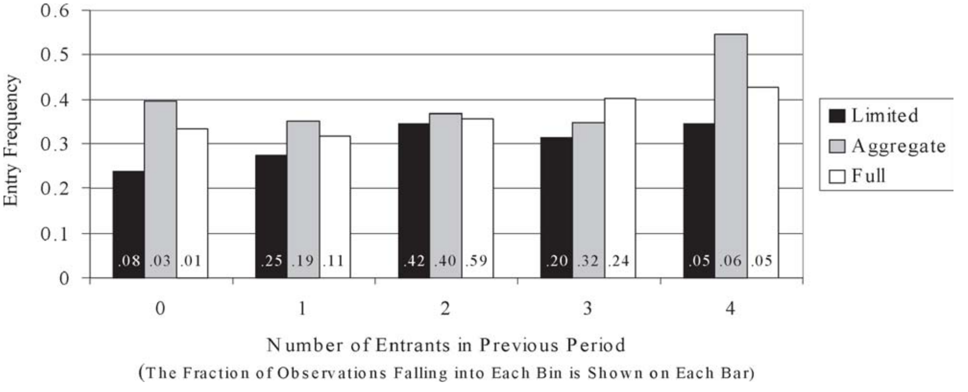


Fig. 3. Frequency of entry conditional on the number of entrants in the previous period.

Figure 3 yields two interesting observations. Consider first the entry frequencies conditional on 0 or 1 entrants in the previous period. In the aggregate and full information treatments, where subjects are aware of the past number of entrants, profitable opportunities from entering do not go unexploited; the entry frequencies in period n conditional on 0 or 1 entrants in period $n-1$ all lie between 30–40%. By contrast, in the limited information treatment where subjects were not aware of the number of entrants in period $n-1$, the entry frequencies conditional on 0 or 1 entrants in period $n-1$ are less than 30%. This finding suggests that individuals in the aggregate and full information treatments are indeed conditioning on the additional information they are given concerning the number of entrants. A second observation is that the conditional entry frequencies for the aggregate information treatment are, with one exception (the number of entrants the previous period was 3) greater than the conditional entry frequencies for the full information treatment. Furthermore, the variance in the conditional entry frequencies is lowest for the full information treatment and highest for the aggregate information treatment. One can interpret these findings as suggesting that subjects are conditioning on the additional information they receive in the full information treatment about precisely which players are entering and which are staying out when making their entry decisions.

The information in Fig. 3 is further disaggregated in Fig. 4, which shows the frequency with which players who *entered* in period $n-1$ also chose to enter in period n, conditional on the total number of entrants in period $n-1$.[18] Here we see that for three of the four bins, the frequency of repeated entry is greatest in the full information treatment. One explanation for this finding is that players in the full information treatment seek to establish a reputation as entrants, capitalizing on the fact that the identity of the players who enter is revealed in this treatment in contrast to the other two treatments where the actions chosen by individual players are not revealed. We will return to the issue of repeated game strategies a little later in the paper. An alternative and complementary explanation is that players learn the pure equilibrium more quickly in the full information treatment so the frequency of repeated entry is greater.

[18] The case of 0 entrants in period $n-1$ is therefore excluded.

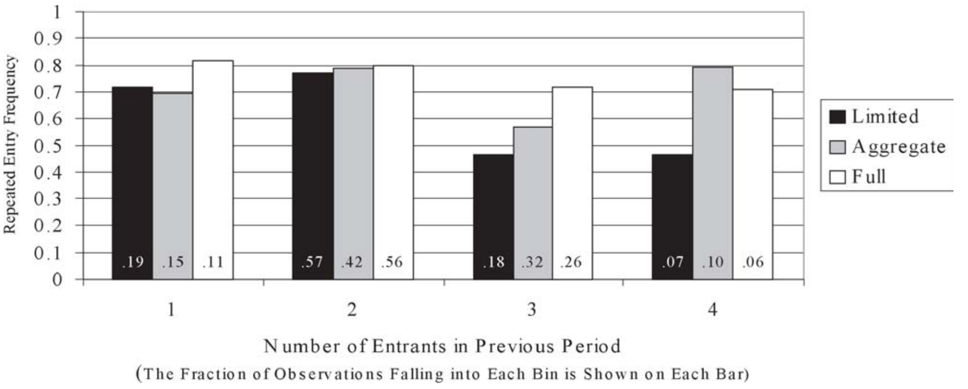


Fig. 4. Frequency of repeated entry conditional on the number of entrants in the previous period.

In a further effort to verify that subjects are responding to the information they are given and to also address the question of whether subjects can be regarded as reinforcement or hypothetical reinforcement learners, we conducted a number of conditional (fixed effects) logit regressions where the dependent variable is the action chosen by subject i in period n, $a_n^i \in \{0, 1\}$, where 1 denotes entry. These logit regressions are of the form:

$$\Pr\bigl[a_n^i = 1\bigr] = \frac{\exp(\alpha^i + \beta_1 O_n^i + \beta_2 H_n^i)}{1 + \exp(\alpha^i + \beta_1 O_n^i + \beta_2 H_n^i)}. \tag{10}$$

Here, α^i is an individual fixed effect specific to player i, O_n^i is individual i's *own* marginal payoff from entry at the start of period n defined by

$$O_n^i = r\sum_{j=1}^{n-1}\delta_j^i(c - m_j)\Big/\sum_{j=1}^{n-1}\delta_j^i, \tag{11}$$

where δ_j^i is an indicator function equal to 1 if player i entered in period j, and 0 otherwise. Similarly, the variable H_n^i is individual i's *hypothetical* marginal payoff from entry at the start of period n defined by

$$H_n^i = r\sum_{j=1}^{n-1}\bigl(1 - \delta_j^i\bigr)(c - m_j - 1)\Big/\sum_{j=1}^{n-1}\bigl(1 - \delta_j^i\bigr). \tag{12}$$

We estimated this conditional logit regression specification for each of the three treatments using pooled data from three sessions of a given treatment. We purged those observations where there was little variation in the entry decision, specifically, where the frequency of entry was less than 0.05, or greater than 0.95.[19] The regression results are reported in Table 6.

[19] In cases where a player (nearly) always enters or (nearly) always stays out, there is a (near) perfect colinearity between the player's action and the individual fixed effect.

Table 6
Estimates from a conditional logit model of the probability of entry

Treatment:	Limited information			Aggregate information			Full information		
Specification:	1	2	3	1	2	3	1	2	3
O_n	0.743***	0.578***	–	0.081	0.090	–	0.156	0.246*	–
	(0.132)	(0.123)		(0.111)	(0.111)		(0.139)	(0.139)	
H_n	-0.335***	–	-0.166**	0.234**	–	0.239**	0.617***	–	0.640***
	(0.090)		(0.083)	(0.117)		(0.117)	(0.146)		(0.145)
$-\ln L$	689.6	696.9	706.4	750.7	752.7	750.9	655.4	664.6	656.0
L.r. test χ^2		14.57	33.55		4.10	0.54		18.43	1.27
$p > \chi^2$		0.00	0.00		0.04	0.46		0.00	0.26
No. obs.	1386	1386	1386	1386	1386	1386	1287	1287	1287

Note: Standard errors in parentheses.
* Significantly different from zero at the 10% level.
** Idem., 5%.
*** Idem., 1%.

For both the aggregate and full information treatments, we find that subjects are significantly more likely to enter the higher is the marginal *hypothetical* payoff from entry. Subjects' *own* marginal payoff from entry appears not to matter in these two treatments. Indeed, a likelihood ratio test suggests that we cannot reject the null hypothesis of no significant difference between specification 1, which includes both O_n and H_n as regressors and specification 3, which purges O_n as a regressor. We conclude that subjects act as hypothetical reinforcement learners in the environments where hypothetical reinforcement is possible.

In the limited information treatment, we find that specification 1 is preferred to both specifications 2 and 3 which purge one of the two explanatory variables. While we would expect the coefficient on O_n to be significantly positive in the limited information case as indeed it is, contrary to our expectations, we find that the coefficient on H_n is significantly different from zero though it has a negative sign. The significance of H_n in explaining entry decisions in the limited information case may seem puzzling, as subjects in this treatment did not have access to the information necessary to construct this hypothetical payoff variable. The puzzle is easily resolved by noting that H_n and O_n are negatively related; indeed, one can rewrite (12) as

$$H_n^i = \left\{ -O_n^i \sum_{j=1}^{n-1} \delta_j^i + r \left[\sum_{j=1}^{n-1} (c - m_j - 1) + \sum_{j=1}^{n-1} \delta_j^i \right] \right\} \Big/ \sum_{j=1}^{n-1} \left(1 - \delta_j^i\right), \tag{13}$$

so a negative coefficient on H_n may simply reflect the positive association between O_n and the probability of entering. We conclude that in the limited information treatment, it is primarily the subjects' own marginal payoff from entry that is significant in explaining their probability of entering, a conclusion that is consistent with the notion that players are reinforcement learners in this environment.

Our logit model specification assumes that players are playing a sequence of independent one-shot games and are not employing dynamic, repeated game strategies that become possible when players are made aware of the payoff function and their repeated interaction with the same group of players (as in our aggregate and full information treatments). As a check on the reasonableness of this assumption, we searched for evidence that players were employing dynamic, repeated game strategies. We can think of at least two types of dynamic strategies for the market entry game (we recognize there are many possibilities). The most obvious is a collusive strategy, e.g. where each player takes a turn as the sole entrant for a period, that yields payoffs that exceed those obtained in the static equilibria. Tables 2–3, which report the mean payoffs and number of entrants suggest that there is no evidence that players adopted such collusive strategies. A second type of dynamic strategy is a reputation-building or "teaching" one where an individual repeatedly enters without regard to the decisions of others and bears any associated cost so as to build a reputation as an entrant. Such a strategy might be supported by the belief that the short-term costs (e.g. due to excess entry) are more than outweighed by the long-term (equilibrium) gain to being one of two entrants and earning a premium of \$0.20 per round relative to the payoff of non-entrants. To check whether subjects were playing such teaching strategies we examined individual payoffs over time looking for evidence of a long sequence of losses followed by a long sequence of gains. Figure 5 reports the sequence of 10-round average

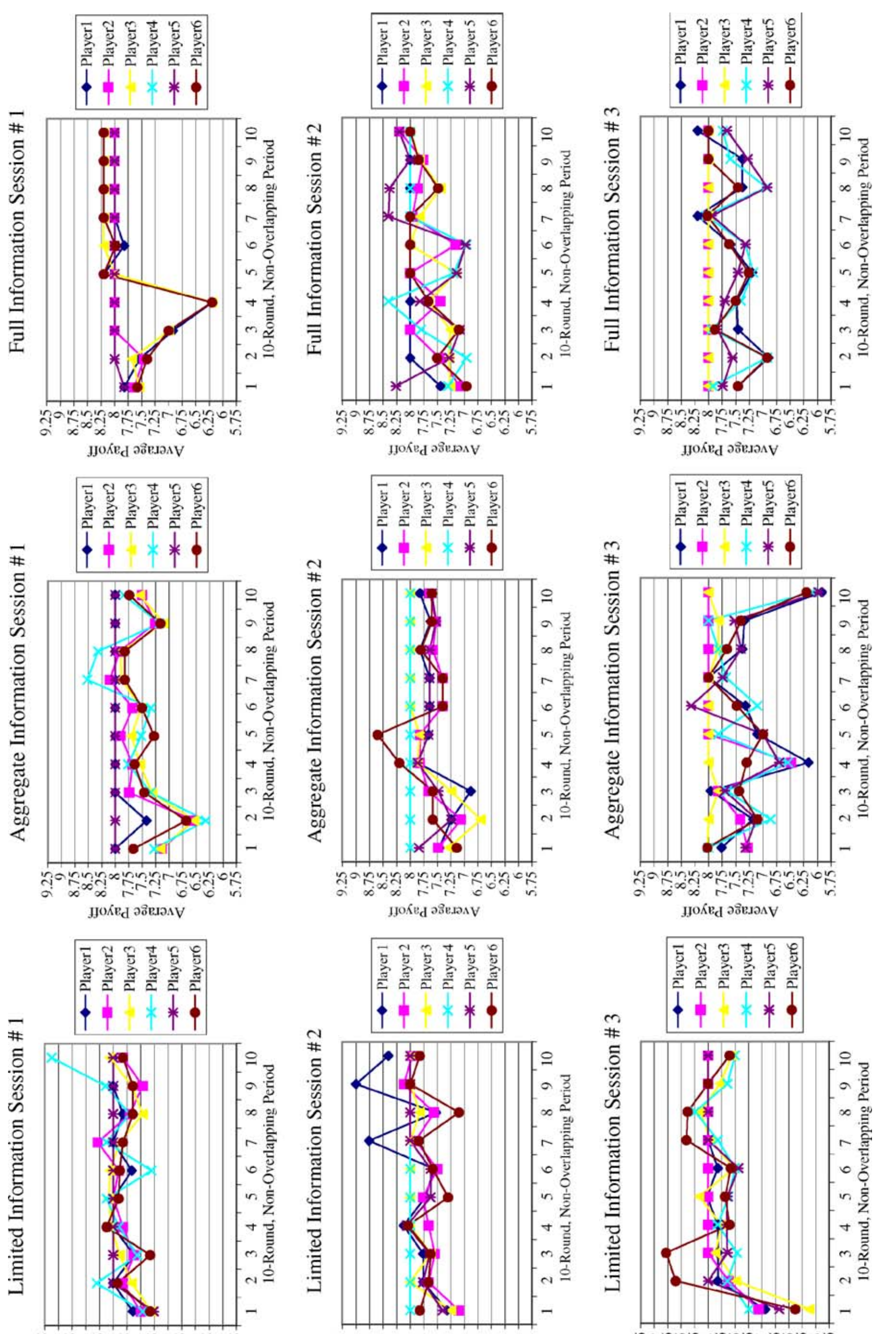

Limited Information Session # 1
Aggregate Information Session # 1
Full Information Session # 1
Limited Information Session # 2
Aggregate Information Session # 2
Full Information Session # 2
Limited Information Session # 3
Aggregate Information Session # 3
Full Information Session # 3
Average Payoff
10-Round, Non-Overlapping Period
Player1
Player2
Player3
Player4
Player5
Player6

Fig. 5.

payoffs for every subject in our experiment. The only instance we found in support of the second type of dynamic strategy described above is in full information session #1 (see the upper rightmost panel of Fig. 5). There we see that three players, numbers 1, 3 and 6, appear to compete for the two entry slots at some cost in terms of average payoff: notice that their average payoff falls to $6.20 over the fourth 10-round period. The other three players always stay out following the first two 10-round periods. Player number 1 eventually stops entering and the pure equilibrium obtains beginning with round 54. No other clear instances were found of this teaching or reputation-building strategy.

Using the 10-round average payoff data for individual subjects presented in Fig. 5, we calculated the autocorrelation in individual payoffs. A negative autocorrelation in payoffs might indicate teaching behavior by that individual. We found that among players who did not always stay out or always enter, the autocorrelation coefficient was significantly different from zero for only three individuals (according to t-statistics for a significance level of 0.05).[20] Of course, it may be that teaching occurs at frequencies other than the 10-round averages that we examined.

We conclude that the logit model specification, which posits that players simply look at round-by-round changes in payoff information, is not unreasonable.

Regarding the individual behavior in full information session #1, the interpretation that players employed dynamic strategies is not inconsistent with our learning hypothesis that players eventually learn the pure strategy equilibrium. Indeed, in this particular session players did learn the pure strategy equilibrium. Learning is, after all, a dynamic process by itself. One could argue that the behavior of the individuals in full information session #1 is not due to their use of dynamic strategies but is instead due to heterogeneity in the belief updating processes. Without resorting to heterogeneity, one can ask whether learning theories, static equilibrium predictions, or repeated game strategies provide the most consistent explanation for the outcomes we observe across all three treatments of our experiment. The evidence we have reported suggests that predictions based on learning theory are the most relevant to understanding our findings.

8. Conclusions

We have derived new results on learning behavior in market entry games and have carried out an experiment to test our predictions. The theoretical predictions appear to have some support. In most sessions, toward the end of 100 rounds, play was at or close to the pure equilibrium outcome predicted by the reinforcement and fictitious play learning models. These findings suggest that it may take a substantial number of repetitions before the play of experimental subjects in market entry games (and possibly other games as well) approaches the asymptotic predictions of learning models. Consequently, caution appears called for in using asymptotic results for learning models to predict or characterize behavior in economic decision-making experiments, which are typically conducted for relatively shorter lengths of time.

[20] It was found to be significantly negative for player #1 of low information session #1, and significantly positive for player #2 of low information session #3 and player # 3 of aggregate information session #2.

Our experimental design also enabled us to investigate subjects' use of information. Our main conclusion here is that individuals are adaptable in ways that are not captured by current learning models. When individuals possess the minimal amount of information assumed by reinforcement learning models, as in our limited information treatment, such that they do not even know that they are playing a game, they are still capable of learning equilibrium behavior. However, reinforcement learning does not capture the change in behavior that occurs when more information is provided. Similarly, belief based learning models, such as fictitious play, do not capture the qualitative difference in play between our aggregate and full information treatments.

One possible explanation for the differences we observe is that individuals are using repeated game (dynamic) strategies that are not captured by the learning models considered. The most common class of repeated game strategies are collusive strategies that permit players to gain greater payoffs than they would in a static equilibrium. There is no evidence for that type of behavior here. We are left to speculate what other objectives the subjects might have had, and what dynamic strategies, out of an infinite class, might have been employed. Identification of these different alternatives is not easy. A second possibility, in line with the work of Camerer et al. (2002), is that certain "sophisticated" players are using the repeated nature of the game and the information about individual actions that is available in the full information treatment to teach other, less sophisticated agents how to play (e.g. to stay out). We found only weak evidence in support of this teaching hypothesis, but perhaps that is because we do not examine strategic behavior across a variety of different repeated games as Camerer et al. (2002) do.

In any case, no single learning model appears to capture the behavior observed across our three experimental treatments. We hope that our analysis has shed some light on the shortcomings of existing learning models, and spurs other researchers to provide further improvements.

Acknowledgments

We thank David Cooper, Jack Ochs and Amnon Rapoport, two referees and an associate editor for helpful comments and suggestions and the US National Science Foundation for research support. Errors remain our own.

Appendix A

This appendix gives the proofs behind the results in the text. We analyze stochastic processes of the form

$$x_{n+1} - x_n = \gamma_n f(x_n) + \gamma_n \eta_n(x_n) + O\left(\gamma_n^2\right) \tag{A.1}$$

for $x_n \in \mathbb{R}^n$. We can think of η as the random component of the process with $E[\eta_n | x_n] = 0$. γ_n is the step size of the process. For all the learning models we consider γ_n is a strictly decreasing sequence, with $\sum_n \gamma_n = \infty$ and $\sum_n \gamma_n^2 < \infty$. This follows from the assumption that the players place an equal weight on every observation.[21]

To obtain results on the asymptotic behavior of these stochastic learning processes, we examine the behavior of the mean or averaged ordinary differential equations (ODEs) derived from the stochastic process above as follows:

$$\dot{x} = f(x). \tag{A.2}$$

We show that in fact the averaged ODEs arising from both reinforcement learning and stochastic fictitious play are both closely related to the evolutionary replicator dynamics (7).

In particular, we apply two classic results from the theory of stochastic approximation. First, Corollary 6.6 of Benaïm (1999) states that if the dynamic (A.2) admits a strict Liapunov function and possesses a finite number of equilibrium points, then with probability one the stochastic process (A.1) must converge to one of these points. We show below that suitable Liapunov functions exist for this class of games for all learning models we consider. Second, Theorem 1 of Pemantle (1990) establishes that the stochastic process (A.1) will converge to an unstable equilibrium point of (A.2) with probability zero. This is important in that we can show that all mixed strategy equilibria in this class of market entry game are unstable under the replicator dynamics (Lemma 1 below). This combined with the application of Corollary 6.6 of Benaïm (1999) implies that for both reinforcement learning and stochastic fictitious play, convergence must be to a pure strategy equilibrium.

First we examine reinforcement learning. Using the results of Hofbauer and Hopkins (2002) it is possible to establish that the mean ODE associated with the model of reinforcement learning given by choice rule (2) and updating rule (4) will be given by the following equations on $[0, 1]^N$:

$$\dot{y}^i = \mu^i y^i \left(1 - y^i\right) r \left(c - 1 - \sum_{j \neq i} y^j\right). \tag{A.3}$$

If each μ^i were exactly one then we would have the standard replicator dynamics. The additional factor μ^i arises because in the original stochastic learning process there is a different step size, equal to $1/Q_n^i$, for each player. We take the step size γ_n of the whole system to be $1/n$, and introduce $\mu^i = n/Q^i > 0$ to keep track of the relative speed of learning of the different players. Because each μ^i is not constant over time, strictly, we also require a further set of equations,

$$\dot{\mu}^i = \mu^i \left(1 - \mu^i \left(v + y^i r \left(c - 1 - \sum_{j \neq 1} y^j\right)\right)\right), \tag{A.4}$$

for $i = 1, 2, \ldots, N$.

[21] There is an alternative hypothesis, for which there is considerable empirical support, that experimental subjects "discount" experience and place greater weight on more recent observations. This would give rise to a constant not decreasing step size. Benaïm and Hirsch (1999) have a result for a class of games that included the current one that if the rate of discount is small then asymptotically play will be close to that generated by learning with a decreasing step size.

Lemma 1. *For market entry games with generic values of c, the only equilibria of the replicator dynamics* (A.3) *together with* (A.4) *which are asymptotically stable are pure Nash equilibria. All other equilibria are unstable.*

Proof. For generic, that is, non integer values of c, this class of market entry games has only a finite number of Nash equilibria each of which is isolated. The fixed points of the replicator dynamics consist of these equilibria and in addition all pure strategy profiles. It can be verified that any equilibrium point of the standard replicator dynamics is an equilibrium point for the joint system (A.3), (A.4).[22]

We first show that the local stability of any such equilibrium is entirely determined by the replicator dynamics and not by the additional equations (A.4). The linearization at any fixed point will be of the form

$$\begin{pmatrix} J & 0 \\ \mathrm{d}\dot{\mu}/\mathrm{d}y & \mathrm{d}\dot{\mu}/\mathrm{d}\mu \end{pmatrix}, \tag{A.5}$$

where J is the Jacobian of the linearized replicator dynamics. Because of the block of zeros to the upper right, it can be shown that every eigenvalue of a matrix of the above form is an eigenvalue for either J or $\mathrm{d}\dot{\mu}/\mathrm{d}\mu$. The latter matrix is diagonal and has only negative elements. Hence, if J has one or more positive eigenvalues, the equilibrium point is unstable for the joint dynamics, if it has N negative eigenvalues, the equilibrium point is asymptotically stable for the joint dynamics.

We now investigate the structure of J. At any fully mixed equilibrium where all players enter with probability $\bar{y}$, the Jacobian J of the linearized replicator dynamics has the form $J_{ii} = \mu^i(1-2y^i)r(c-1-\sum_{j\neq i} y^j)$ which equals zero if $y^i = \bar{y}$ for $i = 1, \dots, N$. That is, J has a zero trace. The off-diagonal elements will be $J_{ij} = -\mu^i \bar{y}(1-\bar{y})r$. Now, as all players earn the same payoff in a mixed equilibrium, therefore $\mu^i = \mu^j$ for all i, j and J will be symmetric. Thus, it has no complex eigenvalues, and with a zero trace, these real eigenvalues sum to zero. Hence, we have a saddlepoint.

At any asymmetric mixed equilibrium let the first $N-j-k$ players randomize over entry and the remaining $j+k$ players play pure. Then one can calculate that in this case that the Jacobian evaluated at this equilibrium has the form

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

where A is a $(N-j-k)\times(N-j-k)$ block of the form found at a symmetric mixed equilibrium as described above, and C is a diagonal matrix. It is easy to show that the eigenvalues of J consist of the eigenvalues of C, which are negative, and of A, which by the above argument are a mixture of positive and negative.

At any pure profile, one can calculate that the Jacobian is diagonal. Furthermore, if this profile is not a Nash equilibrium then at least one diagonal element is positive. In con-

[22] In fact, for each equilibrium point of the standard replicator dynamics, there are two for the joint system, one with μ positive and the other with μ equal to zero. However, the latter is always unstable and is never an asymptotic outcome for the reinforcement learning process.

trast, at a pure Nash equilibrium which must be strict as c is non-integer, all elements are negative. □

Proof of Proposition 1. As outlined above, the proof is in two steps. We identify a suitable Liapunov function which ensures convergence of the stochastic process. Then, we show that the stochastic process cannot converge to a mixed equilibrium. First, define

$$V_0(y) = r\sum_{i=1}^{N} y^i\left(c-1-\frac{1}{2}\sum_{j\neq i} y^j\right). \tag{A.6}$$

Note that

$$\frac{\partial V_0}{\partial y^i} = c-1-\sum_{j\neq i} y^j.$$

This function has a local maximum at each pure Nash equilibrium and a local minimum at each pure state which is not Nash.

$$\dot{V}_0(y) = \frac{\mathrm{d}V_0(y)}{\mathrm{d}y}\cdot\dot{y} = \sum_{i=1}^{N}\mu^i y^i\left(1-y^i\right)\left(r\left(c-1-\sum_{j\neq i}^{j}\right)\right)^2 \geqslant 0$$

with equality only where $\dot{y}=0$. Hence, $V_0(y)$ is a strict Liapunov function in the sense of Corollary 6.6 of Benaïm (1999). Second, for generic values of c, this class of game possesses a finite number of equilibria. Hence, by that Corollary, the stochastic process must converge to an equilibrium point. It is shown in Hopkins and Posch (2002) that this form of reinforcement learning converges to unstable fixed points of the replicator dynamics with probability zero. Hence, play must converge to a pure equilibrium. □

Proof of Proposition 2. In the case of the exponential version of stochastic fictitious play, given the expected motion (8), (see Hopkins, 2002 for details), the associated ODE will be

$$\dot{y}^i = \beta\left(y^i\left(1-y^i\right)r\left(c-1-\sum_{j\neq i} y^j\right) + \frac{1}{\beta}y^i\left(1-y^i\right)\left(\log\left(1-y^i\right)-\log y^i\right)\right). \tag{A.7}$$

Now consider the modified Liapunov function

$$V_1(y) = V_0(y) - \frac{1}{\beta}\sum_{i=1}^{N}\left(y^i\log y^i + \left(1-y^i\right)\log\left(1-y^i\right)\right).$$

Note that

$$\frac{\partial V_1(y)}{\partial y^i} = r\left(c-1-\sum_{j\neq i} y^j\right) + \frac{1}{\beta}\left(\log\left(1-y^i\right)-\log y^i\right).$$

This implies that, first, the critical points of V_1 correspond to perturbed equilibria of the dynamics (A.7), and second,

$$\dot{V}_1(y) = \frac{\mathrm{d}V_1(y)}{\mathrm{d}y} \cdot \dot{y}$$
$$= \sum_{i=1}^{N} y_n^i \bigl(1 - y_n^i\bigr) \left(r \left(c - 1 - \sum_{j \neq i} y_n^j \right) + \frac{1}{\beta} \bigl(\log\bigl(1 - y^i\bigr) - \log y^i\bigr) \right)^2 \geqslant 0$$

with equality only where $\dot{y} = 0$. Hence, $V_1(y)$ is a strict Liapunov function in the sense of Corollary 6.6 of Benaïm (1999). Second, for generic values of c, this class of game possesses a finite number of equilibria. Hence, by that Corollary, the stochastic process must converge to an equilibrium point. With the exponential dynamics (A.7), as β becomes large, the dynamics approach a positive scalar transformation of the replicator dynamics (7). So for β large enough the results of Lemma 1 will hold. Therefore, by Theorem 1 of Pemantle (1990), convergence to any equilibrium other than a pure Nash equilibrium is impossible. □

References

Anderson, C.M., Camerer, C.F., 2000. Experience-weighted attraction learning in sender–receiver signaling games. Econ. Theory 16, 689–718.

Benaïm, M., 1999. Dynamics of stochastic algorithms. In: Azéma, J., et al. (Eds.), Séminaire de Probabilités XXXIII. In: Lecture Notes in Mathematics, vol. 1709. Springer-Verlag, Berlin, pp. 1–68.

Benaïm, M., Hirsch, M.W., 1999. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. Games Econ. Behav. 29, 36–72.

Camerer, C.F., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67, 827–874.

Camerer, C.F., Lovallo, D., 1999. Overconfidence and excess entry: an experimental approach. Amer. Econ. Rev. 89, 306–318.

Camerer, C.F., Ho, T.-H., Chong, J.K., 2002. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. J. Econ. Theory 104, 137–188.

Erev, I., Barron, G., 2002. On adaptation, maximization and reinforcement learning among cognitive strategies. Working paper. Technion, Haifa.

Erev, I., Rapoport, A., 1998. Coordination, "magic," and reinforcement learning in a market entry game. Games Econ. Behav. 23, 146–175.

Erev, I., Roth, A.E., 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. Amer. Econ. Rev. 88, 848–881.

Fudenberg, D., Levine, D.K., 1998. The Theory of Learning in Games. MIT Press, Cambridge, MA.

Hofbauer, J., Hopkins, E., 2002. Learning in perturbed asymmetric games. Working paper. University of Edinburgh. Available from http://homepages.ed.ac.uk/ehk/.

Hopkins, E., 2002. Two competing models of how people learn in games. Econometrica 70, 2141–2166.

Hopkins, E., Posch, M., 2002. Attainability of boundary points under reinforcement learning. Working paper. Univ. of Edinburgh. Available from http://homepages.ed.ac.uk/ehk/.

McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6–38.

Monderer, D., Shapley, L.S., 1996. Fictitious play property for games with identical interests. J. Econ. Theory 68, 258–265.

Ochs, J., 1998. Coordination in market entry games. In: Budescu, D.V., Erev, I., Zwick, R. (Eds.), Games and Human Behavior. Erlbaum, Mahwah, NJ, pp. 143–172.

Pemantle, R., 1990. Nonconvergence to unstable points in urn models and stochastic approximations. Ann. Probability 18, 698–712.

Rapoport, A., Seale, D.A., Erev, I., Sundali, J.A., 1998. Equilibrium play in large group market entry games. Manage. Sci. 44, 119–141.

Rapoport, A., Seale, D.A., Winter, E., 2000. An experimental study of coordination and learning in iterated two-market entry games. Econ. Theory 16, 661–687.
Rapoport, A., Seale, D.A., Winter, E., 2002. Coordination and learning behavior in large groups with asymmetric players. Games Econ. Behav. 39, 137–166.
Sarin, R., Vahid, F., 1999. Payoff assessments without probabilities: a simple dynamic model of choice. Games Econ. Behav. 28, 294–309.
Seale, D.A., Rapoport, A., 2000. Elicitation of strategy profiles in large group coordination games. Exper. Econ. 3, 153–179.
Sundali, J.A., Rapoport, A., Seale, D.A., 1995. Coordination in market entry games with symmetric players. Organ. Behav. Human Dec. Process. 64, 203–218.