IJSM, Volume 11, Number 1, 2011 ISSN: 1555-2411

INTERNATIONAL JOURNAL OF STRATEGIC MANAGEMENT®

SPONSORED BY:

Angelo State University
San Angelo, Texas, USA
www.angelo.edu

Managing Editors:

Professor Alan S. Khade, Ph.D.
California State University Stanislaus

Professor Detelin Elenkov, Ph.D. (MIT)
Angelo State University, Texas, USA

A Publication of the
International Academy of Business and Economics®
IABE.EU

LEARNING AN OPPONENT'S STRATEGY IN COURNOT COMPETITION

C.-Y. Cynthia Lin, University of California at Davis, Davis, USA

ABSTRACT

This paper analyzes the dynamics of learning to compete strategically in a Cournot duopoly. The learning in games model used is logistic smooth fictitious play. I develop novel software that can be used to confirm and visualize existing analytic results, to generate ideas for future analytic proofs, to analyze games for which analytic solutions are difficult to derive, and to aid in the teaching of learning in games in a graduate game theory, business strategy, or business economics course. One key result is that there is an overconfidence premium: the worse off a player initially expects her opponent to be, the better off she herself will eventually be.

Keywords: *stochastic fictitious play, learning in games*

1. INTRODUCTION

Although most work in non-cooperative game theory has traditionally focused on equilibrium concepts such as Nash equilibrium and their refinements such as perfection, models of learning in games are important for several reasons. The first reason why learning models are important is that mere introspection is an insufficient explanation for when and why one might expect the observed play in a game to correspond to an equilibrium. For example, experimental studies show that human subjects often do not play equilibrium strategies the first time they play a game, nor does their play necessarily converge to the Nash equilibrium even after repeatedly playing the same game (see e.g., Erev & Roth, 1998). In contrast to traditional models of equilibrium, learning models appear to be more consistent with experimental evidence (Fudenberg & Levine, 1999). These models, which explain equilibrium as the long-run outcome of a process in which less than fully rational players grope for optimality over time, are thus potentially more accurate depictions of actual real-world strategic behavior. By incorporating exogenous common shocks, this paper brings these learning theories even closer to reality.

In addition to better explaining actual strategic behavior, the second reason why learning models are important is that they can be useful for simplifying computations in empirical work. Even if they are played, equilibria can be difficult to derive analytically and computationally in real-world games. For cases in which the learning dynamics converge to an equilibrium, deriving the equilibrium from the learning model may be computationally less burdensome than attempting to solve for the equilibrium directly. Indeed, the fictitious play learning model was first introduced as a method of computing Nash equilibria (Hofbauer & Sandholm, 2001). More recently, Pakes and McGuire (2001) use a model of reinforcement learning to reduce the computational burden of calculating a single-agent value function in their algorithm for computing symmetric Markov perfect equilibria. As will be explained below, the work presented in this paper further enhances the applicability of these learning models to empirical work.

In this paper, I use one commonly used learning model: stochastic fictitious play. I analyze the dynamics of a particular form of stochastic fictitious play—logistic smooth fictitious play—and apply my analysis to a Cournot duopoly. I analyze the following issues, among others:

- (i) Trajectories*:* What do the trajectories for strategies, assessments, and payoffs look like?
- (ii) Convergence: Do the trajectories converge? Do they converge to the Nash equilibrium? How long does convergence take?
- (iii) Welfare: How do payoffs from stochastic fictitious play compare with those from the Nash equilibrium? When do players do better? Worse?
- (iv) Priors: How do the answers to (i)-(iii) vary when the priors are varied?

I develop novel software that can be used to confirm and visualize existing analytic results, to generate ideas for future analytic proofs, to analyze games for which analytic solutions are difficult to derive, and to aid in the teaching of learning in games in a graduate game theory, business strategy, or business economics course.

My analyses yield several central results. First, varying the priors affects the distribution of production and of payoffs between the two firms, but not either the weighted sum of expected quantity produced nor the weighted sum of payoff achieved. Second, there is an *overconfidence premium*: the worse off a player initially expects her opponent to be, the better off she herself will eventually be.

The balance of this paper proceeds as follows. I describe my model in Section 2. I outline my methods and describe my software in Section 3. In Section 4, I analyze the Cournot duopoly dynamics in the benchmark case without Nature. Section 5 concludes.

2. MODEL

2.1 Cournot duopoly

The game analyzed in this paper is a static homogeneous-good Cournot duopoly. I choose a Cournot model because it is one of the widely used concepts in applied industrial organization (Huck, Normann & Oeschssler, 1999); Although the particular game I analyze in this paper is a Cournot duopoly, my software can be used to analyze any static normal-form two-player game.I focus on two firms only so that the phase diagrams for the best response dynamics can be displayed graphically. Although my software can only generate phase diagrams for two-player games, it can be easily modified to generate other graphics for games with more than two players.

In a one-shot Cournot game, each player *i* chooses a quantity q_i to produce in order to maximize her one-period profit (or payoff):

$$\pi_i(q_i, q_j) = D^{-1}(q_i + q_j)q_i - C_i(q_i)$$

where $D^{-1}(\bullet)$ is the inverse market demand function and $C_i(q_i)$ is the cost to firm *i* of producing q_i. Each firm *i*'s profit-maximization problem yields the best-response function:

$$BR_i(q_j) = \arg\max_{q_i} \pi_i(q_i, q_j)$$

I assume that the market demand $D(\bullet)$ for the homogeneous good as a function of price p is linear and is given by:

$$D(p) = a - bp$$

where $a \geq 0$ and $b \geq 0$. I assume that the cost $C_i(\bullet)$ to each firm *i* of producing q_i is quadratic and is given by:

$$C_i(q_i) = c_i q_i^2$$

where $c_i \geq 0$. With these assumptions, the one-period payoff to each player i is given by:

$$\pi_i(q_i, q_j) = \left(\frac{a}{b} - \frac{q_i + q_j}{b}\right)q_i - c_i q_i^2 \quad ,$$

the best-response function for each player *i* is given by:

$$BR_i(q_j) = \frac{(a - q_j)}{2(1 + c_i b)}$$

and the Nash equilibrium quantity for each player *i* is given by:

$$q_i = \frac{a(1 + 2c_j b)}{4(1 + c_i b)(1 + c_j b) - 1} \quad .$$

Throughout the simulations, I set a = 20, b = 1. With these parameters, the maximum total production $\bar{q}$, corresponding to p = 0, is $\bar{q} = 20$. The pure-strategy space S^i for each player i is thus the set of integer quantities from 0 to $\bar{q} = 20$. I examine two cases in terms of cost functions. In the symmetric case, I set $c_1 = c_2 = 1/2$; in the asymmetric case, the higher-cost player 1 has $c_1 = 4/3$, while the lower-cost player 2 has $c_2 = 5/12$. The Nash equilibrium quantities are thus q_1^{NE} = 5, q_2^{NE} = 5 in the symmetric case and q_1^{NE} = 3, q_2^{NE} = 6 in the asymmetric case. These correspond to payoffs of π_1^{NE} = 37.5, π_2^{NE} = 37.5 in the symmetric case and π_1^{NE} = 21, π_2^{NE} = 51 in the asymmetric case. The monopoly profit or, equivalently, the maximum joint profit that could be achieved if the firms cooperated, is π^m = 80 in the symmetric case and π^m =75.67 in the asymmetric case.

As a robustness check, I also run all the simulations under an alternative set of cost parameters. The alternative set of parameters in the symmetric cost case are $c_1 = c_2 = 0$, which yields a Nash equilibrium quantity of q_1^{NE} = q_2^{NE} = 5 and a Nash equilibrium payoff of π_1^{NE} = π_2^{NE} = 37.5. The alternative set of parameters in the asymmetric cost case are $c_1 = 0.5$, $c_2 = 0$, which yields Nash equilibrium quantities of q_1^{NE} = 4, q_2^{NE} = 8 and Nash equilibrium payoffs of π_1^{NE} = 24, π_2^{NE} = 64. Except where noted, the results across the two sets of parameters have similar qualitative features.

2.2 Logistic smooth fictitious play

The one-shot Cournot game described above is played repeatedly and the players attempt to learn about their opponents over time. The learning model I implement is that of stochastic fictitious play. In fictitious play, agents behave as if they are facing a stationary but unknown distribution of their opponents' strategies; in stochastic fictitious play, players randomize when they are nearly indifferent between several choices (Fudenberg & Levine, 1999). The particular stochastic play procedure I implement is that of logistic smooth fictitious play.

Although the one-shot Cournot game is played repeatedly, I assume, as is standard in learning models, that current play does not influence future play, and therefore ignore collusion and other repeated game considerations. As a consequence, the players regard each period-t game as an independent one-shot Cournot game. There are several possible stories for why it might be reasonable to abstract from repeated play considerations in this duopoly setting. One oft-used justification is that each period there is an anonymous random matching of the firms from a large population of firms (Fudenberg & Levine, 1999). This matching process might represent, for example, random entry and/or exit behavior of firms. It might also depict a series of one-time markets, such as auctions, the participants of which differ randomly market by market. A second possible story is that legal and regulatory factors may preclude collusion.

For my model of logistic smooth fictitious play, I use notation similar to that used in Fudenberg and Levine (1999). As explained above, the pure-strategy space S^i for each player i is the set of integer quantities from 0 to $\bar{q} = 20$. A pure strategy q_i is thus an element of this set: $q_i \in S^i$. The per-period payoff to each player i is simply the profit function $\pi_i(q_i, q_j)$.

At each period t, each player i has an assessment $\gamma_t^i(q_j)$ of the probability that his opponent will play q_j. This assessment is given by

$$\gamma_t^i(\tilde{q}_j) = \frac{\kappa_t^i(\tilde{q}_j)}{\sum_{q_j=0}^{\bar{q}} \kappa_t^i(q_j)},$$

where the weight function $\kappa_t^i(q_j)$ is given by:

$$\kappa_t^i(q_j) = \kappa_{t-1}^i(q_j) + \mathrm{I}\{q_{j,t-1} = q_j\}$$

with exogenous initial weight function $\kappa_0^i(q_j) : S^j \to \Re_+$. Thus, for all periods *t*, γ_t^i, κ_t^i and κ_0^i are all $1 \times \bar{q}$ vectors. For my various simulations, I hold the length of the fictitious history, $\sum_{q_j=0}^{\bar{q}} \kappa_t^i(q_j)$, constant at $\bar{q} + 1 = 21$ and vary the distribution of the initial weights.

In logistic smooth fictitious play, at each period *t*, given her assessment $\gamma_t^i(q_j)$ of her opponent's play, each player *i* chooses a mixed strategy σ_i so as to maximize her perturbed utility function:

$$\tilde{U}^i(\sigma_i, \gamma_t^{\ i}) = E_{q_i,q_j}[\pi_i(q_i, q_j) \mid \sigma_i, \gamma_t^{\ i}] + \lambda v^i(\sigma_i),$$

where $v^i(\sigma_i)$ is an admissible perturbation of the following form:

$$v^i(\sigma_i) = \sum_{q_i=0}^{\bar{q}} -\sigma_i(q_i) \ln \sigma_i(q_i).$$

With these functional form assumptions, the best-response distribution $\overline{BR^i}$ is given by

$$\overline{BR^i}(\gamma_t^i)[\tilde{q}_i] = \frac{\exp(1/\lambda)\,\mathrm{E}_{q_i}[\pi_i(\tilde{q}_i, q_j) \mid \gamma_t^i]}{\sum_{q_i=0}^{\bar{q}} \exp(1/\lambda)\,\mathrm{E}_{q_i}[\pi_i(q_i, q_j) \mid \gamma_t^i]}.$$

The mixed strategy θ_t^i played by player *i* at time *t* is therefore given by:

$$\theta_t^i = \overline{BR^i}(\gamma_t^i) \quad .$$

The pure action q_{it} actually played by player *i* at time *t* is drawn for player *i*'s mixed strategy:

$$q_{it} \sim \theta_t^{\ i}.$$

Because each of the stories of learning in static duopoly I outlined above suggest that each firm only observes the play of its opponent and not the plays of other firms of the opponent's "type" in identical and simultaneous markets, I assume that each firm only observes the actual pure-strategy action q_{it} played by its one opponent and not the mixed strategy θ_t^i from which that play was drawn.

I choose the logistic model of stochastic fictitious play because of its computational simplicity and because it corresponds to the logit decision model widely used in empirical work (Fudenberg & Levine, 1999). For the simulations, I set $\lambda = 1$.

3. METHODS AND SOFTWARE

To analyze the dynamics of logistic smooth fictitious play, novel software is developed that enables one to analyze the following.

(i) Trajectories

For each player i, I examine the trajectories over time for the mixed strategies θ_t^i chosen, the actual pure actions q_{it} played and payoffs π_{it} achieved. I also examine, for each player i, the trajectories for the per-period mean quantity of each player's mixed strategy:

$$\mathrm{E}[q_i \mid \theta_t^i] \qquad (1)$$

as well as the trajectories for the per-period mean quantity of his opponent's assessment of his strategy:

$$\mathrm{E}[q_i \mid \gamma_t^j]\,. \qquad (2)$$

I also examine three measures of the players' payoffs. I examine the payoffs (or, equivalently, profits) instead of the perturbed utility so that I can compare the payoff from stochastic fictitious play with the payoffs from equilibrium play. First, I examine the *ex ante payoffs*, which I define to be the payoffs a player expects to achieve before her pure-strategy action has been drawn from her mixed strategy distribution:

$$\mathrm{E}_{q_i,q_j}[\pi_i(q_i,q_j) \mid \theta_t^i, \gamma_t^i]\,. \qquad (3)$$

The second form of payoffs are the *interim payoffs*, which I define to be the payoffs a player expects to achieve after she knows which particular pure-strategy action q_{it} has been drawn from her mixed strategy distribution, but before her opponent has played:

$$\mathrm{E}_{q_j}[\pi_i(q_{it},q_j) \mid \gamma_t^i] \qquad (4)$$

The third measure of payoffs I analyze is the actual realized payoff $\pi_i(q_{it}, q_{jt})$.

(ii) Convergence

The metric I use to examine convergence is the Euclidean norm $d(\bullet)$. Using the notion of a Cauchy sequence and the result that in finite-dimensional Euclidean space, every Cauchy sequence converges (Rudin, 1976), I say that a vector-valued trajectory $\{X_t\}$ has converged at time τ if for all $m,n \geq \tau$ the Euclidean distance between its value at periods m and n, $d(X_m, X_n)$, falls below some threshold value $\bar{d}$. In practice, I set $\bar{d} = 0.01$ and require that $d(X_m, X_n) < \bar{d}\ \forall m,n \in [\tau, T]$, where T=1000. I examine the convergence of two trajectories: the mixed strategies $\{\theta_t^i\}$ and ex ante payoffs $\{\mathrm{E}_{q_i,q_j}[\pi_i(q_i,q_j) \mid \theta_t^i, \gamma_t^i]\}$.

In addition to analyzing whether or not either the mixed strategies or the ex ante payoffs converge, I also examine whether or not they converge to the Nash equilibrium strategy and payoffs, respectively. I say that a vector-valued trajectory $\{X_t\}$ has converged to the Nash equilibrium at time τ if the Euclidean distance between its value at and that of the Nash equilibrium analog, $d(X_t, X^{NE})$, falls below some threshold value $\bar{d}$ for all periods after τ. In practice, I set $\bar{d} = .01$ and require that $d(X_t, X^{NE}) < \bar{d} \quad \forall t \in [\tau, T]$, where T=1000.

(iii) Welfare

The results above are compared to non-cooperative Nash equilibrium as well as the cooperative outcome that would arise if the firms acted to maximize joint profits. The cooperative outcome corresponds to the monopoly outcome.

(iv) Priors

Finally, I examine the effect of varying both the mean and spread of players' priors κ_0, the above results. These priors reflect the initial beliefs each player has about his opponent prior to the start of play.

The software developed for analyzing the dynamics of logistic smooth fictitious play can be used for several important purposes. First, this software enables one to confirm and visualize existing analytic results. For example, for classes of games for which convergence results have already been proven, my software enables one not only to confirm the convergence, but also to visualize the transitional dynamics. I demonstrate such a use of the software in my analysis of a Cournot duopoly.

A second way in which my software can be used is to generate ideas for future analytic proofs. Patterns gleaned from computer simulations can suggest results that might then be proven analytically. For example, one candidate for an analytic proof is the result that, when costs are asymmetric and priors are uniformly weighted, the higher-cost player does better under stochastic fictitious play than she would under the Nash equilibrium. Another candidate is the result is what I term the *overconfidence premium*: the worse off a player initially expects her opponent to be, the better off she herself will eventually be.

A third way in which of my software can be used is to analyze games for which analytic solutions are difficult to derive.

A fourth potential use for my software is pedagogical. The software can supplement standard texts and papers as a learning or teaching tool in any course covering learning dynamics and stochastic fictitious play.

I apply the software to analyze the stochastic fictitious play dynamics of the Cournot duopoly. Although the entire software was run for two sets of parameters, I present the results from only one. Unless otherwise indicated, qualitative results are robust across the two sets of parameters.

4. RESULTS

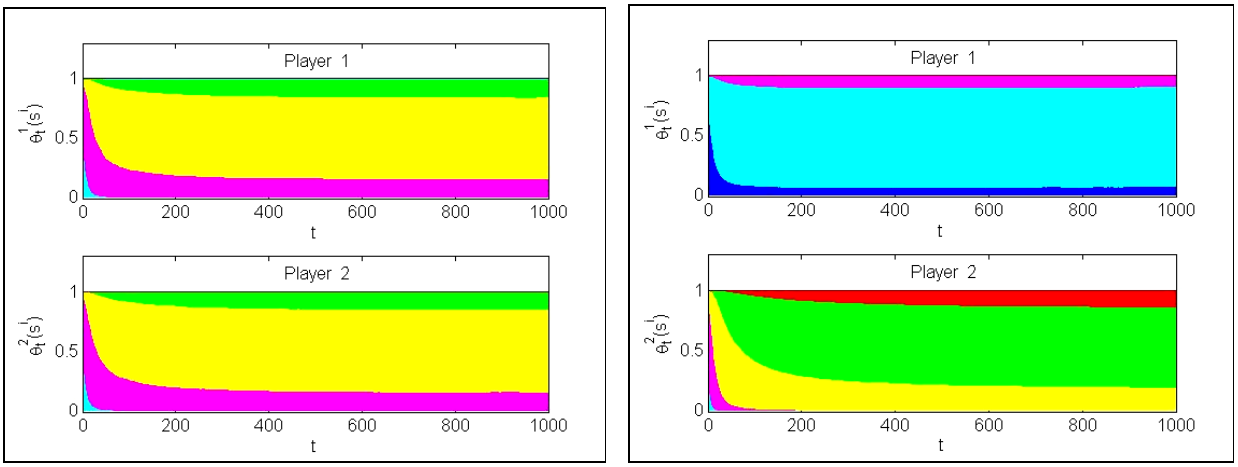
I analyze the stochastic fictitious play dynamics of the Cournot duopoly game. Because my game is a 2X2 game that has a unique strict Nash equilibrium, the unique intersection of the smoothed best response functions is a global attractor (Fudenberg & Levine, 1999). Since my Cournot duopoly game with linear demand therefore falls into a class of games for which theorems about convergence have already been proven, a presentation of my results enables one not only to confirm the previous proven analytic results, but also to assess how my numerical results may provide additional information and intuition previously inaccessible to analytic analysis alone.

First, I present results that arise when each player initially believes that the other plays each possible pure strategy with equal probability. In this case, each player's prior puts uniform weight on all the possible pure strategies: κ_0^i=(1,1, …, 1) $\forall i$. I call this form of prior a *uniformly weighted prior*. When a player has a uniformly weighted prior, he will expect his opponent to produce quantity 10 on average, which is higher than the symmetric Nash equilibrium quantity of q_1^{NE} = q_2^{NE} = 5 in the symmetric cost case and also higher than both quantities q_1^{NE} = 3, q_2^{NE} = 6 that arise in the Nash equilibrium of the asymmetric cost case.

Figure 1 presents the trajectories of each player *i*'s mixed strategy θ_t^i over time when each player has a uniformly weighted prior. Each color in the figure represents a pure strategy (quantity) and the height of the band represents the probability of playing that strategy. As expected, in the symmetric case, the players end up playing identical mixed strategies. In the asymmetric case, player 1, whose costs are higher, produces smaller quantities than player 2. In both cases the players converge to a fixed mixed strategy, with most of the change occurring in the first 100 time steps. It seems that convergence takes longer in the case of asymmetric costs than in the case of symmetric costs. Note that the strategies that eventually dominate each player's mixed strategy initially have very low probabilities. The explanation for this is that with uniformly weighted priors, each player is grossly overestimating how much the other will produce. Each player expects the other to produce quantities between 0 and 20 with equal probabilities, and thus has a mean prior of quantity 10. As a consequence, each firm initially produces much less the

Nash equilibrium quantity to avoid flooding the market. In subsequent periods, the players will update their assessments with these lower quantities and change their strategies accordingly.

FIGURE 1

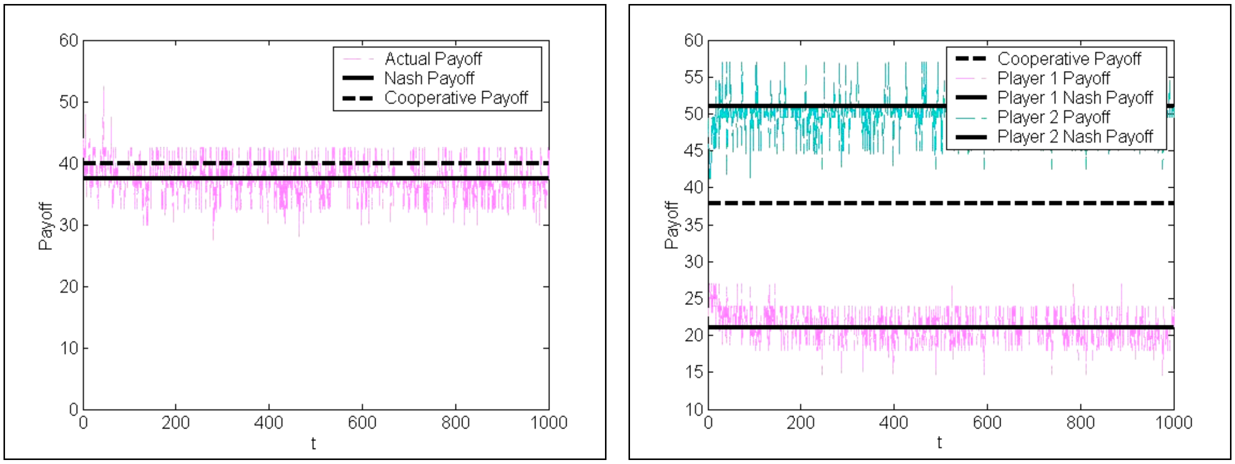


(a) (b)

Dynamics of players' mixed strategies with (a) symmetric and (b) asymmetric costs as a function of time. As a benchmark, the Nash equilibrium quantities are $q^{NE} = (5,5)$ *in the symmetric cost case and* $q^{NE} = (3,6)$ *in the asymmetric cost case. Each player has a uniformly weighted prior.*

Figure 2 presents the trajectories for the actual payoffs π_{it} achieved by each player *i* at each time period *t*. Once again, I assume that each player has a uniformly weighted prior. The large variation from period to period is a result of players' randomly selecting one strategy to play from their mixed strategy vectors. In the symmetric case, each player *i*'s per-period payoff hovers close to the symmetric Nash equilibrium payoff of π_i^{NE} = 37.5. On average, however, both players do slightly worse than the Nash equilibrium, both averaging payoffs of 37.3 (s.d. = 2.96 for player 1 and s.d. = 2.87 for player 2). The average and standard deviation for the payoffs are calculated as follows: means and standard deviations are first taken for all *T*=1000 time periods for one simulation, and then the values of the means and standard deviations are averaged over 20 simulations. In the asymmetric case, the vector of players' per-period payoffs is once again close to the Nash equilibrium payoff vector π^{NE} = (21, 51). However, player 1 slightly outperforms her Nash equilibrium, averaging a payoff of 21.16 (s.d. = 2.16), while player 2 underperforms, averaging a payoff of 50.34 (s.d. = 2.59). Thus, when costs are asymmetric, the high-cost firm does better on average under logistic smooth fictitious play than in the Nash equilibrium, while the low-cost firm does worse on average. This qualitative result is robust across the two sets of cost parameters I analyzed.

FIGURE 2



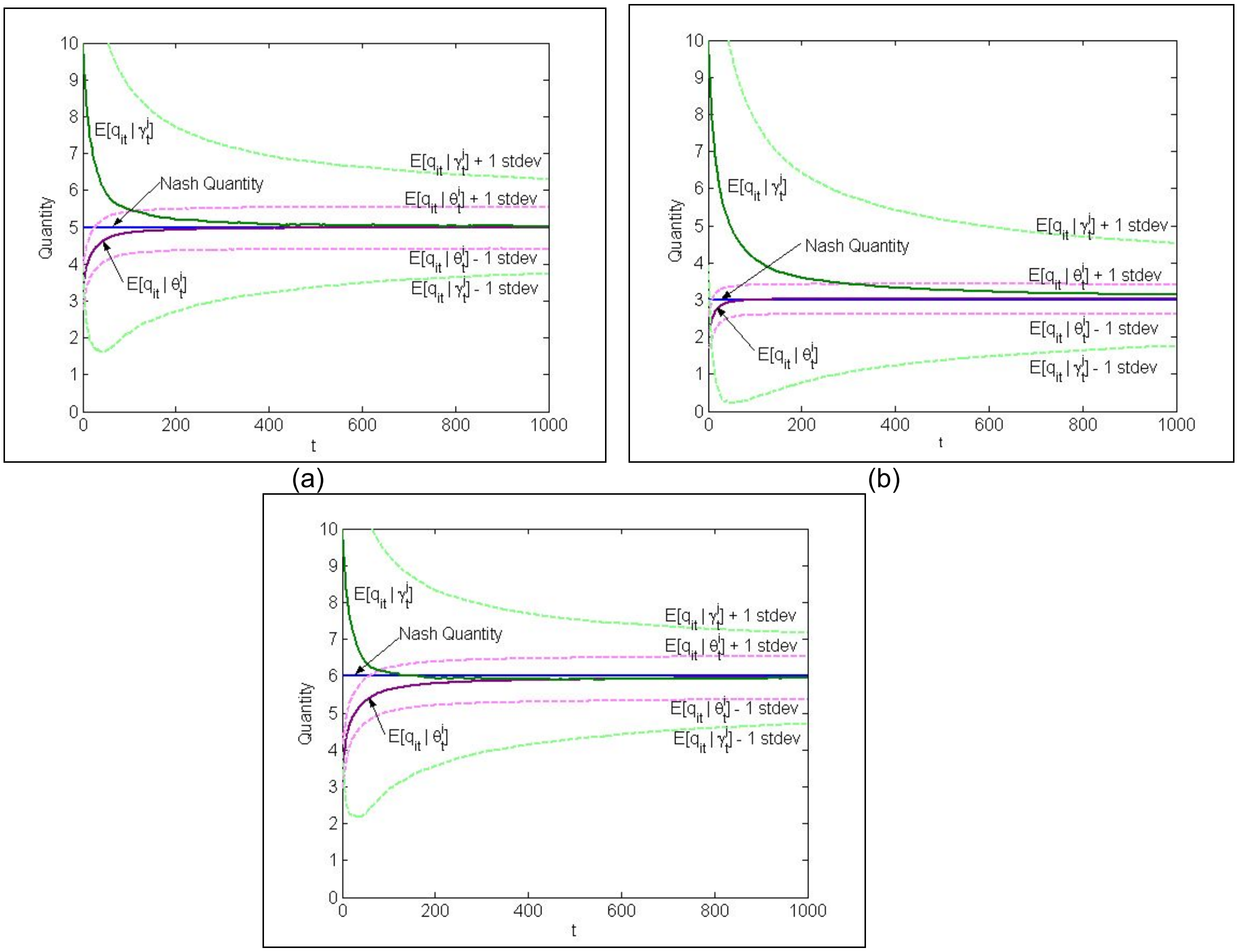
(a) (b)

Actual payoffs achieved by each player as a function of time in the (a) symmetric and (b) asymmetric cases. Each player has a uniformly weighted prior.

Much of the variation in the achieved payoff arises from the fact at each time *t*, each player *i* randomly selects one strategy q_{it} to play from his time-*t* mixed strategy vector θ_t^i. By taking the mean over these vectors at each time *t*, I can eliminate this variation and gain a clearer picture of the dynamics of each player's strategy. Figure 3 presents the evolution of the expected per-period quantities, where expectations are taken at each time *t* either over players' mixed strategies or over opponents' assessments at time *t*, values corresponding to expressions (1) and (2), respectively. As before, each player has a uniformly weighted prior. Figures 3(a) and 3(b) present the both mean of player 1's mixed strategy (i.e., $\mathrm{E}[q_1 \mid \theta_t^1]$) and the mean of player 2's assessment of what player 1 will play (i.e., $\mathrm{E}[q_i \mid \gamma_t^j]$) for the symmetric- and asymmetric-cost cases, respectively. Figure 3(c) gives the mean of player 2's mixed strategy and the mean of player 1's assessment of player 2 in the asymmetric case.

For both the symmetric and asymmetric cost cases, the mean of player 2's assessment is initially very high and asymptotically approaches the Nash equilibrium. As explained above, this is a result of the uniformly weighted prior. Initially, player 2 expects player 1 to play an average strategy of 10. Similarly, player 1 expects player 2 to play an average strategy of 10, and consequently player 1's mixed strategy initially has a very low mean, which rises asymptotically to the Nash equilibrium. It is interesting to note that in the asymmetric case, the mean over player 1's chosen mixed strategy slightly overshoots the Nash equilibrium and then trends back down towards it. Figure 3 also provides standard deviations over player 1's mixed strategy and player 2's assessment. Note that the standard deviation of player 2's assessment is much higher than the standard deviation of player 1's mixed strategy, indicating player 2's relative uncertainty about what player 1 is doing. Although the results presented in these figures are the outcome of one particular simulation, in general the variation in the values for the expected quantities across simulations is small.

FIGURE 3



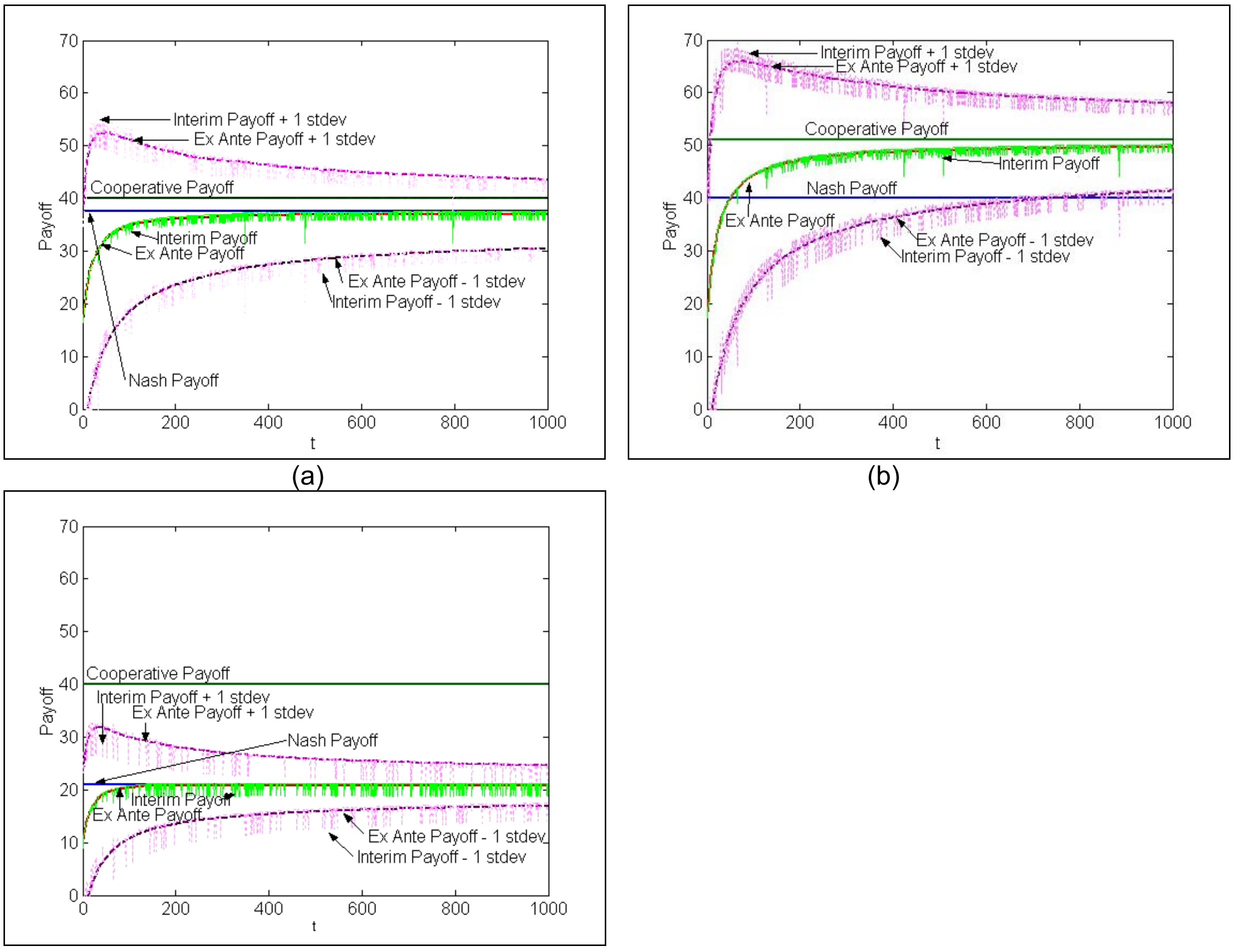
(a) (b)

(c)

Means and variances of quantities, as taken over players' time-t mixed strategies and opponent's time-t assessments in the (a) symmetric case and the asymmetric case for (b) the higher-cost player 1 and (c) the lower-cost player 2 as a function of time. Each player has a uniformly weighted prior.

Just as an examination of the expected per-period quantity instead of the mixed strategy vector can elucidate some of the dynamics underlying play, analyzing expected payoffs can similarly eliminate some of the variation present in the trajectories of players' achieved payoffs in Figure 2. Figure 4 presents the evolutions of players' ex ante and interim expected payoffs, corresponding to expressions (3) and (4), respectively. Figures 4(a) and 4(b) depict these quantities for player 1. The interim payoff has a large variance from period to period because it is calculated after player 1 has randomly selected a strategy from his mixed strategy. In the symmetric case, depicted in Figure 4(a), both the ex ante and interim expected payoffs asymptote to the Nash equilibrium payoff, but remain slightly below it. In the asymmetric cost case, the high-cost player 1 eventually does better than she would have in Nash equilibrium while the low-cost player 2 eventually achieves approximately his Nash equilibrium payoff. In the alternative set of cost parameters I tried, the high-cost player 1 eventually achieves approximately her Nash equilibrium payoff in the long run while the low-cost player 2 does worse than his Nash equilibrium payoff. For all cases, on average, the interim expected payoff is below the ex ante expected payoff. Figure 4 also presents standard deviations for the ex ante and interim expected payoffs; in general they seem roughly equal. As before, although the results presented in these figures are the outcome of one particular simulation, in general the variation in the values for the ex ante and interim payoffs across simulations is small. Thus, while players in the symmetric cost case do slightly worse than in Nash equilibrium in the long run, the high-cost player 1 in the asymmetric cost case does better in the long run under stochastic fictitious play than she would in Nash equilibrium.

FIGURE 4



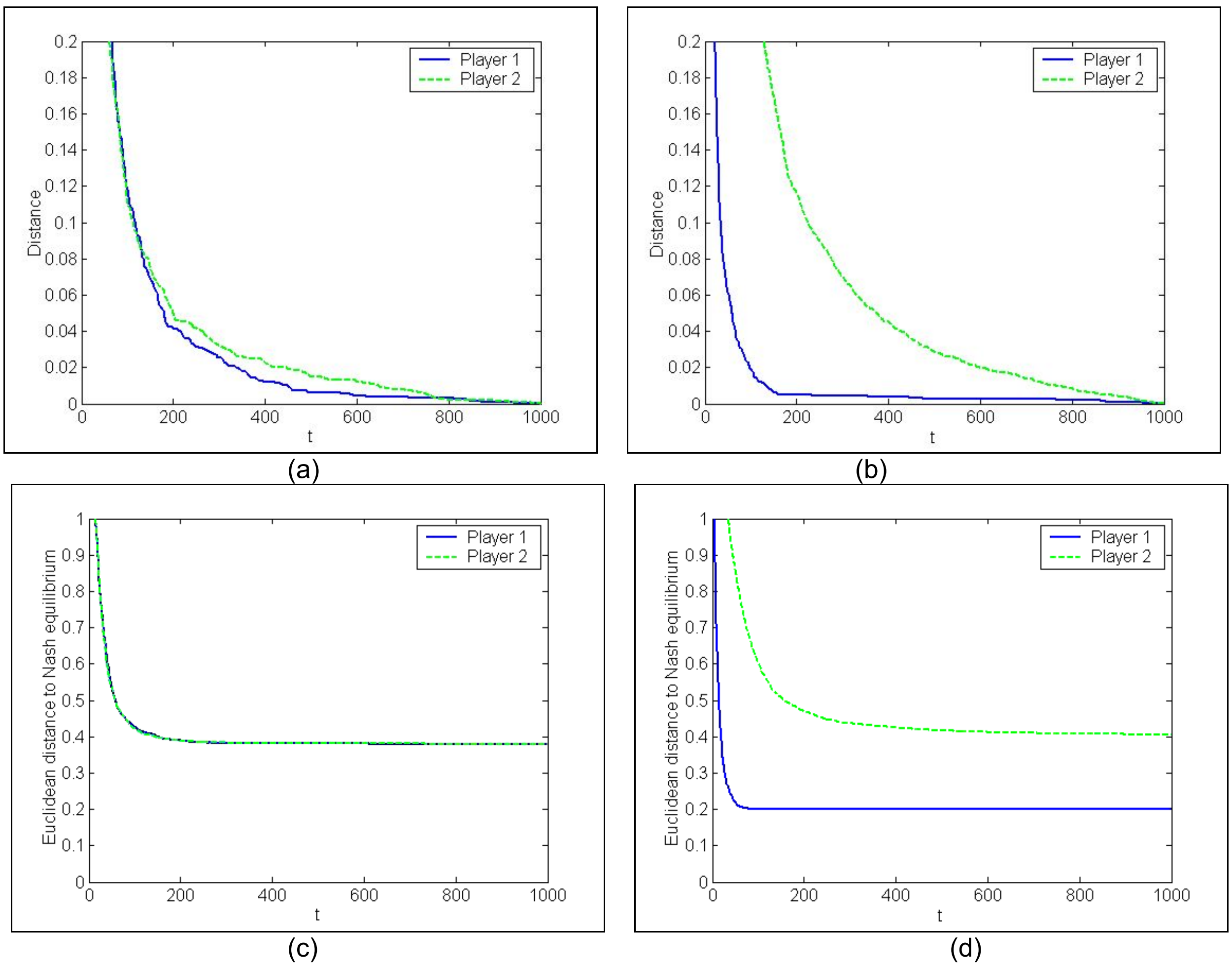
(c)

Means and variances of ex ante and interim payoff in the (a) symmetric case and the asymmetric case for (b) the higher-cost player 1 and (c) the lower-cost player 2 as a function of time. Each player has a uniformly weighted prior.

Having shown that expected quantity and expected payoff seem to converge to the Nash equilibrium, I now test whether this is indeed the case. First, I examine whether or not the mixed strategies do converge and the speed at which they converge. Figure 5 gives a measure of the convergence of smooth fictitious play when priors are uniformly weighted. As explained above, I define how close to steady-state player *i* is at time *t* as the maximum Euclidean distance between player *i*'s mixed strategy vector θ_t^i at times *m, n* ≥ *t*. Indeed, the mixed strategies do converge: the Euclidean distance asymptotes to zero. In the symmetric case, Figure 5(a), both players converge at approximately the same rate. In the asymmetric case, Figure 5(b), the player with higher costs, player 1, appears to converge more quickly.

Now that I have established that the mixed strategies do indeed converge, the next question I hope to answer is whether they converge to the Nash equilibrium. Figures 5(c) and 5(d) depict the Euclidean distance between player i's mixed strategy vector θ_t^i and the Nash equilibrium. In the symmetric case, both players converge at about the same rate, but neither gets very close to the Nash equilibrium. In the asymmetric case, player 1 again stabilizes more quickly. Furthermore, player 1 comes much closer to the Nash equilibrium than player 2 does. With uniformly-weighted priors, it is never the case that $d(X_t, X^{NE}) < \bar{d}$, where $\bar{d} = 0.01$; thus neither player converges to the Nash equilibrium. At time T=1000, the distance to the Nash equilibrium is 0.21 in the symmetric case. In the asymmetric case, the distance to Nash equilibrium is 0.39 for the higher cost player and 0.53 for the lower cost player.

FIGURE 5

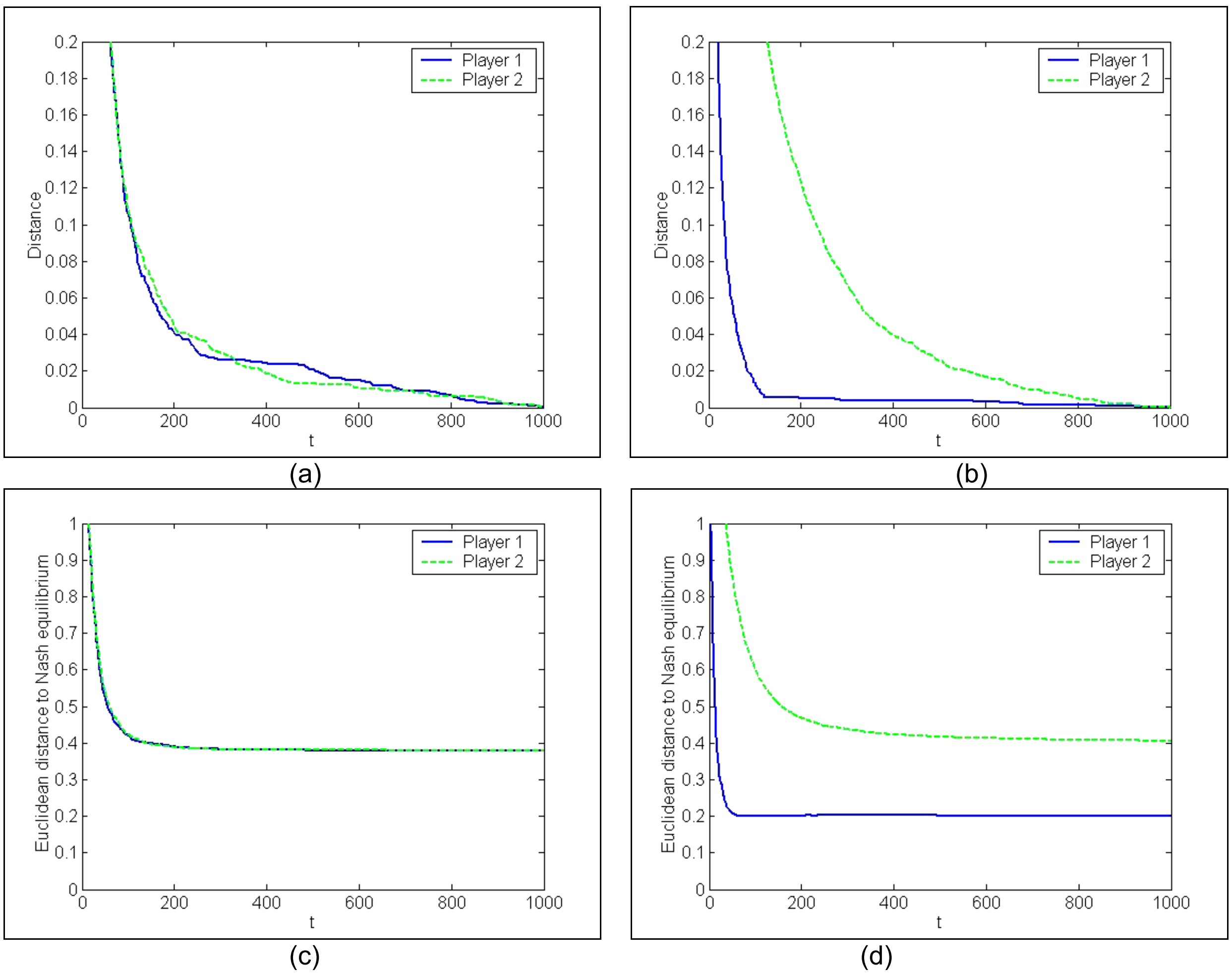


Maximum Euclidean distance between player i's mixed strategy vector θ_t^i in periods m, n ≥ t in the (a) symmetric and (b) asymmetric cases as a function of time. Distance between player i's mixed strategy vector and the Nash equilibrium in the (c) symmetric and (b) asymmetric cases. Each player has a uniformly weighted prior.

Because the players' prior beliefs are responsible for much of the behavior observed in the early rounds of play, I now examine how the mean and the spread of the priors affect the convergence properties. First, I examine how my results may change if instead of a uniformly weighted prior, each player i's had a prior that concentrated all the weight on a single strategy: κ_0^i = (0, 0, …, 21, 0, 0, …, 0). I call such a prior a *concentrated prior*.

Figure 6 repeats the analyses in Figure 5, but with concentrated priors that place all the weight on quantity 9. The figures show that in both the symmetric and asymmetric cases, the form of the prior affects the speed of convergence but not its asymptotic behavior. Even with concentrated priors, each player's play still converges to a steady state mixed strategy vector. With concentrated priors, just as with uniformly weighted priors, the distance to the Nash equilibrium converges to 0.21 in the symmetric case and 0.39 and 0.53 in the asymmetric case.

FIGURE 6

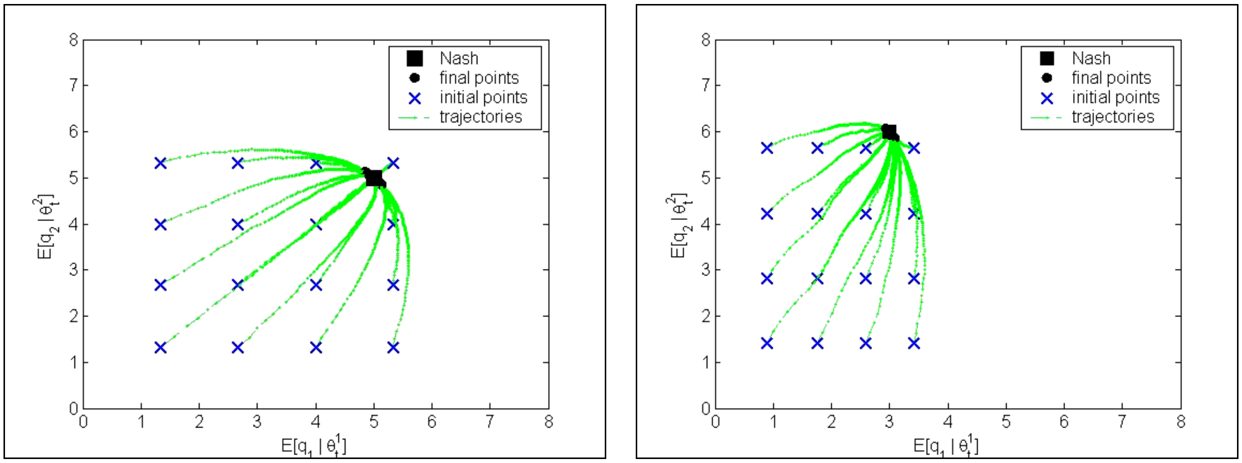


Maximum Euclidean distance between player i's mixed strategy vector θ_t^i in periods m, n ≥ t in the (a) symmetric and (b) asymmetric cases as a function of time. Distance between player i's mixed strategy vector and the Nash equilibrium in the (c) symmetric and (b) asymmetric cases. Each player i has a concentrated prior that places all the weight on the strategy q_j = 9.

I now examine the effect of varying the means of the concentrated priors on the mean quantity $\mathrm{E}[q_i \mid \theta_t^i]$ of each player *i*'s mixed strategy. For each player, I allow the strategy with the entire weight of 21 to be either 4, 8, 12, or 16. Thus, I have 16 different combinations of initial priors. The phase portraits in Figure 7 are produced as follows. For each of these combinations of priors, I calculate each player *i*'s expected quantity over their mixed strategies $\mathrm{E}[q_1 \mid \theta_t^1]$ and plot this as an ordered pair for each time *t*.

Each trajectory thus corresponds to a different specification of the priors, and displays the evolution of the mixed strategy over T=1000 periods. The figure shows that in both cases, no matter what the prior, the players converge to a point close to the Nash equilibrium. In fact, the endpoints, corresponding to T=1000, appear to fall on a line. It is also interesting to note that many of the trajectories are not straight lines, indicating that players are not taking the most direct route to their endpoints. Notice that in the asymmetric case player 2's quantity never gets very far above her Nash equilibrium quantity.

FIGURE 7

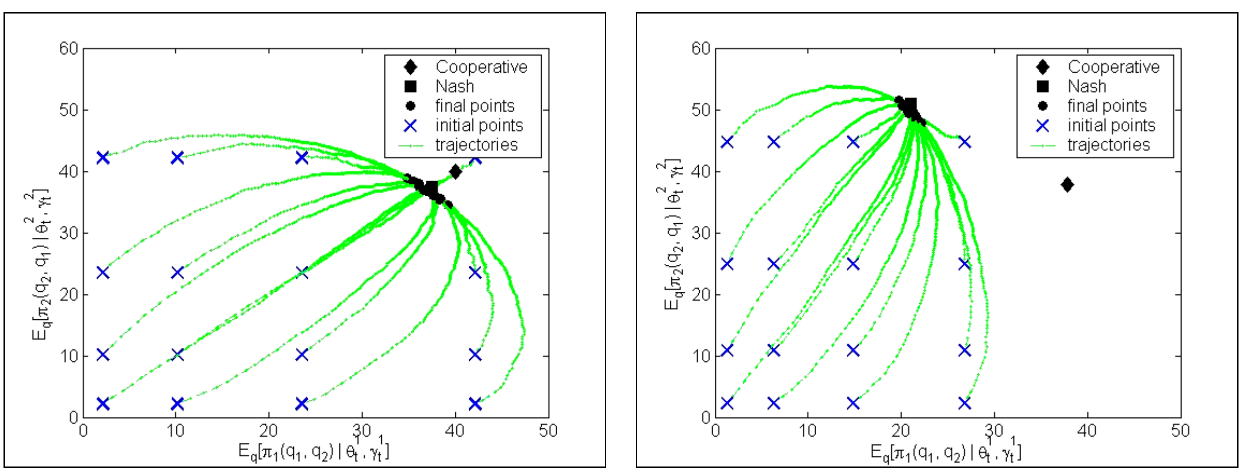


(a) (b)

Phase portraits of expected quantity show the effect of varying (concentrated) priors in the (a) symmetric and (b) asymmetric cases.

While Figure 7 shows phase portraits of expected quantity, Figure 8 shows phase portraits of ex ante expected payoff. For comparison, the payoffs from the Nash and cooperative equilibria are plotted as benchmarks. Once again, no matter the initial prior, the payoffs converge close to the Nash equilibrium payoffs in both cases. Again, the endpoints, corresponding to T=1000, appear to fall on a line. In this case, however, the Nash equilibrium appears to be slightly above the line. Thus, in the steady-state outcome of logistic smooth fictitious play, the players are worse off than they would be in a Nash equilibrium.

FIGURE 8



(a) (b)

Phase portraits of ex ante expected payoff in the (a) symmetric and (b) asymmetric cases.

As noted above, the final points of the trajectories of expected quantity shown in Figure 7 seem to form a line, as do the final points of the trajectories of ex anted expected payoffs in Figure 8. Figure 9 shows only these final points and their best-fit line for both the expected quantity and for the ex ante expected payoff. As seen in Figure 6, each of the final points represents the long-run steady state reached by the players.

Several features of the results in Figure 9 should be noted. The first feature is the linear pattern of the final points. In the symmetric case, the slope of the best-fit line, which lies below the Nash equilibrium, is approximately -1.01 (s.e. = 3e-6). Thus, varying the prior appears only to affect the distribution of production between the two firms, but not the total expected quantity produced, and this total expected quantity is weakly less than that which arises in the Nash equilibrium. In the asymmetric case, the slope of the line, which again lies below the Nash equilibrium, is -1.58 (s.e. = 3e-6).

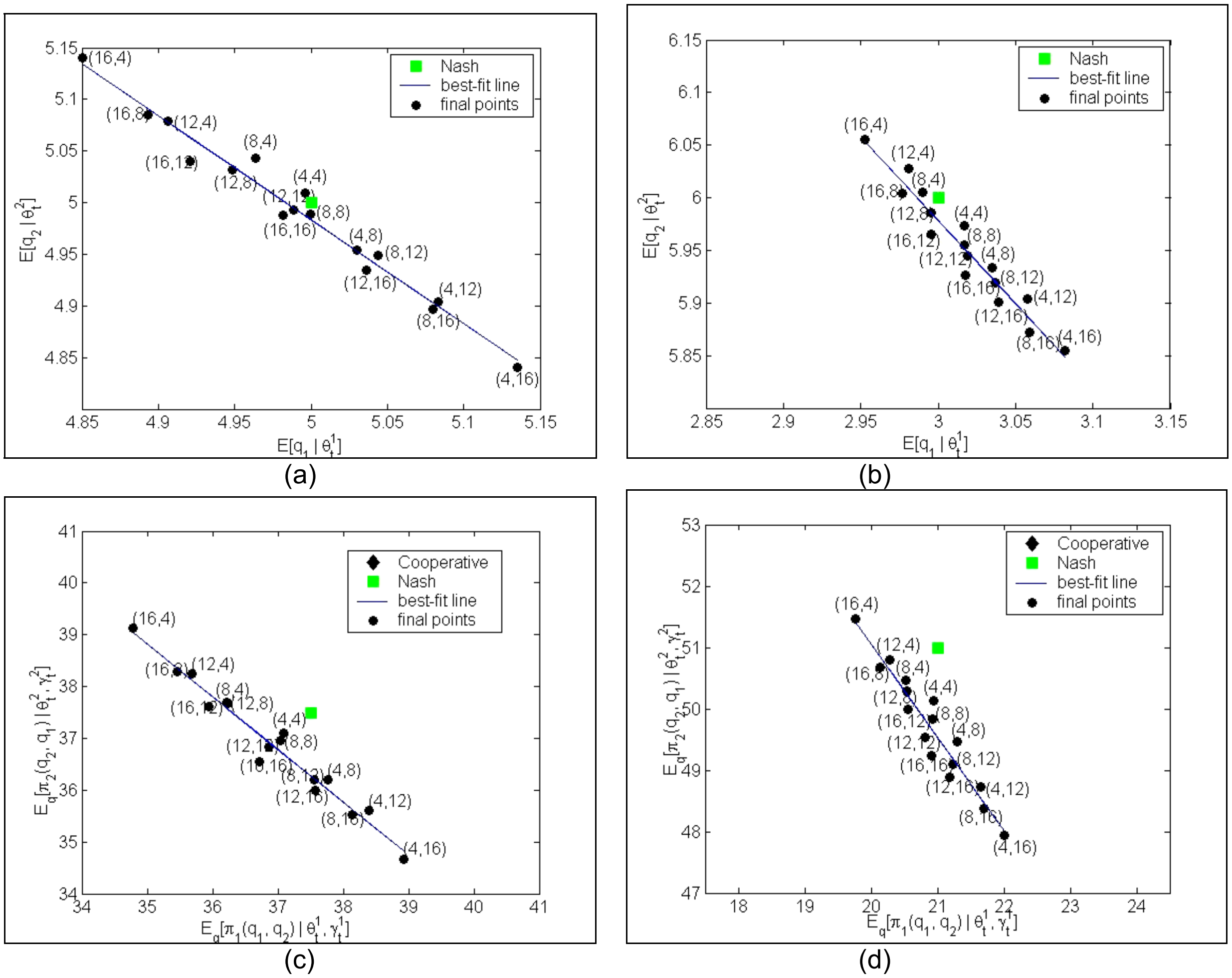
Thus, a weighted sum of the expected quantities, where the higher cost player 1 is given a greater weight, is relatively constant across different priors. Similar statements can be made about the payoffs as well: that is, the sum of the payoffs is robust to the priors but lower than the sum of the Nash equilibrium payoffs in the symmetric cost case, and a weighted sum of the payoffs is robust to the priors but lower than the weighted sum of the Nash equilibrium payoffs in the asymmetric cost case. A second feature of Figure 9 to note regards how each player performs relative to his Nash equilibrium across different priors. In the symmetric case, the final points are distributed fairly evenly about the Nash equilibrium along the best-fit line.

This implies that the number of priors for which player 1 does better than the Nash equilibrium is approximately equal to the number of priors for which player 2 does better than the Nash equilibrium. In the asymmetric case, on the other hand, most of the endpoints lie below the Nash equilibrium, implying that the number of priors for which player 1 does better than the Nash equilibrium is larger than the number of priors for which player 2 does better than the Nash equilibrium.

This seems to confirm the earlier observation that the higher cost player usually outperforms her Nash equilibrium in the asymmetric case. A third important feature of Figure.9 regards convergence. Note that Figures 9(a) and .9(b) show that there are several combinations of priors ((4,4), (8,8), (12, 12), and (16, 16) in the symmetric case, and (8,4), (12, 8) and (16,12) in the asymmetric case) that lead to steady state expected quantities very close to the Nash equilibrium (within a Euclidean distance of 0.01). However, convergence as earlier defined requires that the players' mixed strategy vectors, not the expectations over these vectors, come within a Euclidean distance of 0.01 of the Nash equilibrium. This does not happen for any set of priors; under no combination of priors do the players mixed strategy vectors converge to the Nash equilibrium.

A fourth important feature of Figure 9 regards the effects of a player's prior on his long-run quantity and payoff. According to Figure 9, when the opponent's prior is held fixed, the lower the prior a player has over her opponent (i.e., the less she expects the other to produce), the more she will produce and the higher her per-period profit in the long run. There thus appears to be what I term the *overconfidence premium*: the worse off a player initially expects her opponent to be, the better off she herself will eventually be.

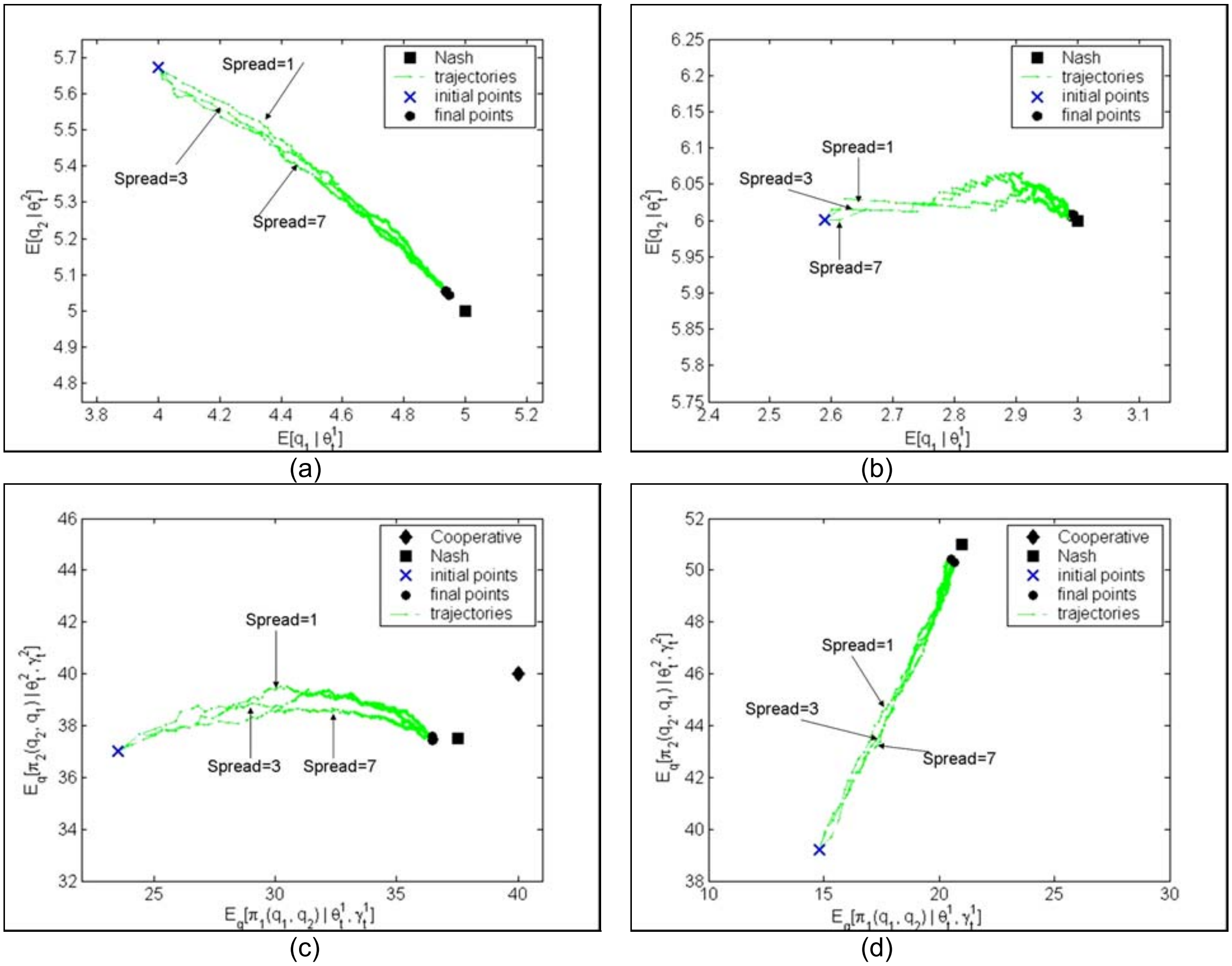
FIGURE 9



Best fit lines of the endpoints of the trajectories of expected quantity shown in Figure 7 in the (a) symmetric and (b) asymmetric cases. Plots (c) and (d) give similar best fit lines for the trajectories of ex ante expected payoff shown in Figure 8.

Having seen the effect of varying the mean of each player's prior on the learning dynamics, I now fix the mean and vary the spread. Figure 10 shows the effect of spread in the prior. I fix player 2's prior, with all weight on one strategy (q_1 = 10). Thus, player 2's prior looks like κ_0^2 = (0, …, 0, 21, 0, … , 0). I chose to concentrate the prior on the mean pure strategy in the strategy set both because it did not correspond to any Nash equilibrium, thus ensuring that the results would be non-trivial, and also so that varying the spread would be straightforward. I vary player 1's prior, keeping its mean the same (also producing a quantity of 10), but spreading its weight over 1, 3, or 7 strategies. Thus, player 1's prior looks like one of κ_0^1 = (0, …, 0, 21, 0, …, 0), (0, …, 0, 7, 7, 7, 0, …0), or (0,…, 0, 3, 3, 3, 3, 3, 3, 3, 0, …, 0). As I see in Figure 10, spreading the prior in this manner does affect the trajectory of expected quantities but does not alter the initial or final points in either the symmetric or asymmetric case. Each trajectory has the exact same starting point, while their final points vary just slightly. This variation decreases as the number of time steps increases. The same result arises if I plot phase portraits of the ex ante payoffs.

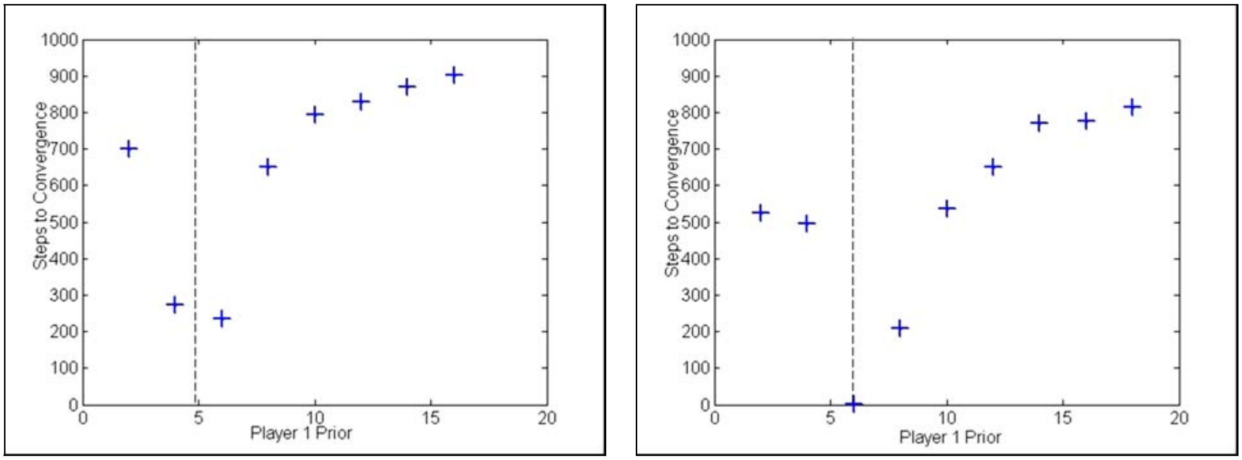
FIGURE 10



The effects of varying the spread of player 1's prior around the same mean (q_2 = 10) on the trajectories of expected quantity in the (a) symmetric and (b) asymmetric cases, and on the trajectories of ex ante payoff in the (c) symmetric and (d) asymmetric cases.

I next examine the effect of varying each player's prior on the rate of convergence. Figure 11 shows the effect of different priors on the speed of convergence of the mixed strategy for player 1. I hold player 2's prior fixed, with all weight on player 1's Nash equilibrium strategy (i.e., $q_1 = q_1^{NE}$ = 5 for the symmetric cost case and $q_1 = q_1^{NE}$ = 3 for the asymmetric cost case). I then vary player 1's prior, keeping all weight on one strategy, but varying that strategy between 2, 4, 6, 8, 10, 12, 14, 16, and 18. Player 2's Nash equilibrium strategy is indicated by the vertical dashed line. In both cases, the time to convergence is minimized when player 1's prior puts all the weight on q_2 = 6. This is not surprising in the asymmetric casE because q^{NE} = (3, 6) is the Nash equilibrium for that case. In the symmetric case, when player 1's prior puts all the weight on q_2 = 6, this is very close to player 2's Nash equilibrium quantity of q_2^{NE} = 5. For the asymmetric cost case, one can also generate an analogous plot as a function of player 2's prior, holding player 1's prior constant at player 2's Nash equilibrium strategy.

FIGURE 11

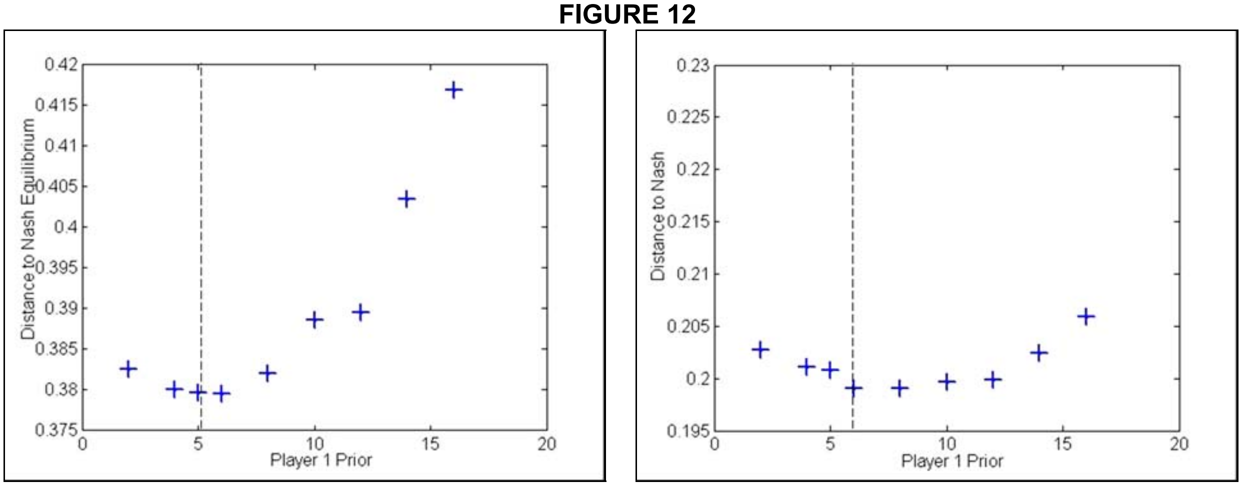


(a) (b)

The number of time steps until convergence in the (a) symmetric and (b) asymmetric cases. Player 1 has a concentrated prior. Player 2's Nash equilibrium strategy is indicated by the dashed line.

I now examine the effect of varying each player's prior on convergence to the Nash equilibrium. Figure 12 shows the effect of different priors on the final distance to the Nash equilibrium. Again, I hold player 2's prior fixed with all weight on player 1's Nash equilibrium strategy. I then vary player 1's prior as before. Player 2's Nash equilibrium strategy is again shown by a dotted vertical line. The distance between player 1's mixed strategy vector and his Nash equilibrium quantity at time T=1000 is smallest when player 1's prior is concentrated at a value close to player 2's Nash equilibrium quantity. The distance grows as player 1's prior gets further away from the Nash equilibrium quantity. For the asymmetric cost case, one can also generate an analogous plot as a function of player 2's prior, holding player 1's prior constant at player 2's Nash equilibrium strategy.

FIGURE 12



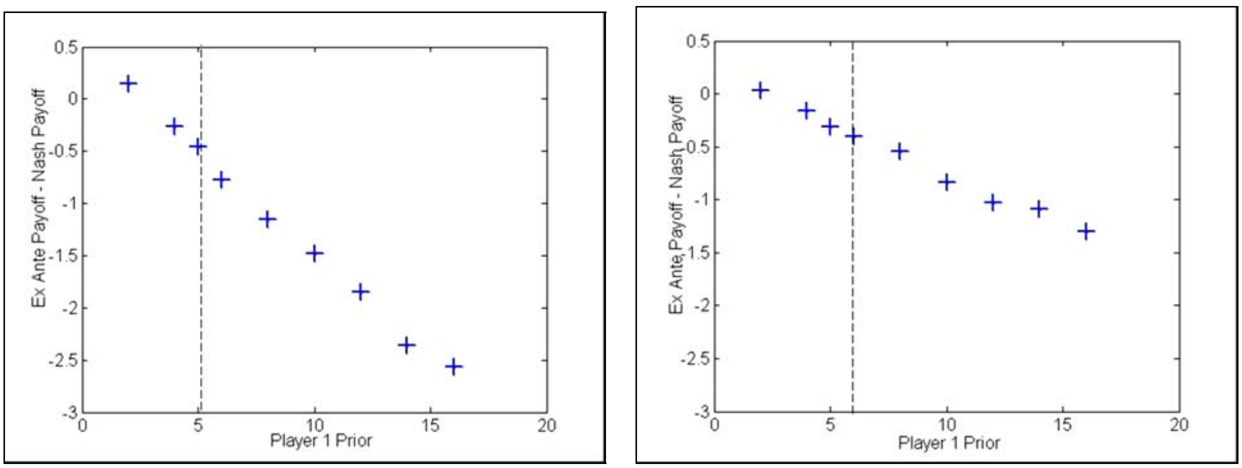
(a) (b)

Distance between player 1's mixed strategy vector and the Nash equilibrium at time T=1000 as a function of player 1's (concentrated) prior in the (a) symmetric and (b) asymmetric cases. Player 2's Nash equilibrium strategy is indicated by the dashed line.

Finally, I examine the effect of varying each player's prior on final-period ex ante payoff, as compared to the Nash equilibrium. Figure 13 shows the effect of different priors on the final ex ante payoff minus the

Nash equilibrium payoff. Again, I hold player 2's prior fixed with all weight on player 1's Nash equilibrium strategy. I then vary player 1's prior as before. Player 2's Nash equilibrium strategy is indicated by a dotted vertical line. The difference between player 1's ex ante payoff and the Nash equilibrium payoff at time T=1000 is largest when player 1's prior is smallest. The difference declines (and becomes negative) as player 1's prior grows. When player 1's prior is smallest, he believes that player 2 will produce a small quantity. Thus, he will produce a large quantity, and reap the benefits of a larger payoff. This result confirms the overconfidence premium results from Figure 9: the worse off a player initially expects his opponent to be, the better off he himself will eventually be. For the asymmetric cost case, one can also generate an analogous plot as a function of player 2's prior, holding player 1's prior constant at player 2's Nash equilibrium strategy.

FIGURE 13



(a) (b)

Difference between player 1's final-period ex ante payoff and the Nash equilibrium payoff at time T=1000 as a function of player 1's concentrated prior in the (a) symmetric and (b) asymmetric cases. Player 2's Nash equilibrium strategy is indicated by the dashed line.

In summary, the main results, which are robust across the two sets of cost parameters analyzed, are:

1) In the symmetric case with uniformly weighted priors, players on average achieve a slightly smaller payoff than the Nash equilibrium payoff, both on average and in the long run.

2) In the asymmetric case with uniformly weighted priors, the higher cost player outperforms her Nash equilibrium payoff both on average and in the long run, while the lower cost player underperforms his on average.

3) With either uniformly weighted priors or concentrated priors, both players' mixed strategy vectors converge to a steady state, but neither player's mixed strategy converges to the Nash equilibrium.

4) In the asymmetric case with uniformly weighted priors, the higher cost player's mixed strategy vector converges to steady state faster than that of the lower cost player. Furthermore, the higher cost player gets closer to the Nash equilibrium.

5) In the symmetric cost case, varying the priors affects the distribution of production and of payoffs between the two firms, but not either the total expected quantity produced nor the total payoff achieved, and both the total quantity and the total payoff are lower than they would be in equilibrium.

6) In the asymmetric cost case, varying the priors affects the distribution of production and of payoffs between the two firms, but not either the weighted sum of expected quantity produced nor the weighted sum of payoff achieved, and both the weighted sum quantity and the weighted sum payoff are lower than they would be in equilibrium.

7) Varying the spread of each player's prior while holding the mean fixed does not affect the long-run dynamics of play.

8) The distance between player 1's mixed strategy vector and his Nash equilibrium quantity at time T=1000 inversely related to the difference between the quantity at which player 1's prior is concentrated and player 2's Nash equilibrium quantity.

9) There is an *overconfidence premium*: the worse off a player initially expects her opponent to be, the better off she herself will eventually be.

These qualitative results are robust across the two sets of cost parameters analyzed.

5. CONCLUSION

In this paper, I investigate a stochastic fictitious play model of learning in a static Cournot duopoly game. Novel software is developed that enables one to analyze the trajectories and convergence properties of strategies, assessments, and payoffs in logistic smooth fictitious play, and to compare the welfare from logistic smooth fictitious play with that from equilibrium play.

My analyses yield several central results. First, varying the priors affects the distribution of production and of payoffs between the two firms, but not either the weighted sum of expected quantity produced nor the weighted sum of payoff achieved. Second, there is an *overconfidence premium*: the worse off a player initially expects her opponent to be, the better off she herself will eventually be. Initial beliefs about the distribution of production and of payoffs can be self-fulfilling.

One key innovation of the work presented in this paper is the novel software, which can be used to confirm and visualize existing analytic results, to generate ideas for future analytic proofs, to analyze games for which analytic solutions are difficult to derive, and to aid in the teaching of stochastic fictitious play in a graduate game theory, business strategy, or business economics course.

REFERENCES:

Erev, I. and Roth, A., "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review*, Volume 88, Pages 848-881, 1998.

Fudenberg, D. and Kreps, D., "Learning Mixed Equilibria", *Games and Economic Behavior*, Volume 5, Pages 320-367, 1993.

Fudenberg, D. and Levine, D., The theory of learning in games, MIT Press, Cambridge, MA, 1999.

Hofbauer, J. and Sandholm, W., "Evolution and Learning in Games with Randomly Disturbed Payoffs", Working paper, Universität Wien and University of Wisconsin, 2001.

Huck, S., Normann, H.-T., and Oeschssler, J. "Learning in Cournot Oligopoly—An Experiment", *Economic Journal*, Volume 109, Number 454, Pages C80-95, 1999.

Li Calzi, M., "Fictitious Play by Cases", *Games and Economic Behavior*, Volume 11, Number 1, Pages 64-89, 1995.

Pakes, A. and McGuire, P., "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality", *Econometrica*, Volume 69, Number 5, Pages 1261-1281, 2001.

Romaldo, D., "Similarities and Evolution", Working paper, 1995.

Rudin, W., Principles of mathematical analysis, McGraw-Hill, Inc, New York,1976.

AUTHOR PROFILE:

Dr. C.-Y. Cynthia Lin is an assistant professor with a joint appointment in the Agricultural and Resource Economics Department and the Environmental Science and Policy Department at the University of California at Davis. She is a member of the Giannini Foundation of Agricultural Economics. Professor Lin received her Ph.D. in Economics from Harvard University in 2006.