Journal of
Economics, Banking and Finance
Volume 2, Issue 1, 2008

Can You Beat A Fictitious Player At Matching Pennies?

Bryan C. McCannon, Assistant Professor of Economics, Wake Forest University, mccannbc@wfu.edu

Abstract:
Fictitious Play is an adaptive play in games that does not require sophisticated players. A fictitious player selects a stationary best response to the historical play of its opponent. While the question of whether games between fictitious players result in convergence to Nash equilibrium has been addressed this work considers the interaction between a fictitious player and a sophisticated player. In the Matching Pennies game optimal play of the sophisticated player is solved for and it is shown that she does better competing against a fictitious player than another sophisticated player. Additionally, the sophisticated player need not act too sophisticatedly. In fact, the sophisticated player can select a simple one-step-ahead plan that brings the empirical frequency of her play closer to the 50-50 threshold in the next stage. Thus, a sophisticated player need not be any more forward-looking than one stage farther than the fictitious player.

Keywords: bounded rationality, competition, Fictitious Play, Matching Pennies

I. Introduction

Adaptive models of play in games attempt to address the critique of standard game theory that it requires unrealistically intelligent, sophisticated individuals who are able to process large amounts of information, compute best responses, update probabilities correctly, and perfectly recall previous actions. Such a critique calls for models that restrict the rationality of the players. An example of an adaptive behavior is Fictitious Play (Brown, 1951; Robinson, 1951) where a player myopically selects a best response to the past play of its opponent using the fraction of rounds where each action was selected as the mixed strategy of the opponent. This particular method has held up reasonably well in experimental work (Boylan and El-Gamal, 1992; Ochs, 1995) and has been used to model learning in games (Fudenberg and Levine, 1998).

Fictitious players make their selection based on the history of play and do not anticipate their opponent's future choices. One might expect this to leave them exposed to manipulation by an intelligent player. This might especially be true in a competitive game such as Matching Pennies. On the other hand Robinson (1951) showed that play between two fictitious players in zero-sum games converges to the equilibrium achieved between two standard players and Miyasawa (1961) illustrated that in all two-player 2x2 games play converges to a Nash equilibrium (see Krishna and Sjöström (1997) for a selective survey of results). Therefore, this note considers the interaction between a sophisticated player and a fictitious player repeatedly playing Matching Pennies to determine whether a sophisticated player can beat a fictitious one.

It is shown that indeed a sophisticated player can beat a fictitious player at Matching Pennies. The game payoff is strictly greater than the renormalized mixed strategy Nash equilibrium payoff received if both players are intelligent. Just how greater this payoff is depends on the fictitious player's choice if indifferent between its two actions. If the probability the fictitious player selects an action is far from the mixed strategy equilibrium, then the sophisticated player's payoff is increased. In fact, if the fictitious player breaks the indifference with a pure action, then the sophisticated player is able to achieve her maximax payoff. A path of play that alternates between the two choices is shown to be optimal. In odd-numbered stages the action that has been selected less often is chosen. This guarantees a win since the fictitious player responds by playing as if the more popular action will be chosen with a higher probability. In even-numbered stages the historical frequency of play of each action is equal and the sophisticated player myopically best responds to the randomization of the fictitious player.

There are some related, theoretical results for play between players who differ in their level of sophistication. Fudenberg and Levine (1995) study universally consistent learning rules; rules that are safe (guarantee minimax payoff) and consistent. Her play is consistent if she does as well as playing a best response to the empirical average of play if the play is given by independent draws from a fixed distribution. Standard fictitious play is consistent, but not safe, and both exponential fictitious play and sophisticated play is universally consistent. They show that in zero-sum games the value of the game is the minimax value if players select universally consistent rules, and they need not necessarily use the same rule. Thus, this holds for competition between a sophisticated player and an exponentially fictitious player. Cahn (2004) studies an adaptive process where agents play according to their regret. Situations where one agent plays according to her regret while others do not play "too sophisticated" are shown to converge to zero regret and correlated equilibrium arise. Vriend (1997) adds intelligence to one agent in two-player ultimatum games and shows that the subgame perfect equilibrium may not arise, but considers reinforcement learners. Ellison (1997) considers a population of players randomly paired to compete in 2x2 coordination games. Only one member of the population is sophisticated, while the rest are fictitious players. He is interested in whether this one sophisticated player acts myopically or is willing to act to shift the population's play from the risk-dominated to the risk-dominant outcome. Thus, the results presented here contribute to the understanding of the effect of asymmetric intelligence in games.

Finally, there is experimental evidence of heterogeneity among a population. Ochs (1995) provides experimental evidence in 2x2 games with a unique mixed strategy equilibrium where the two behaviors most commonly observed were "equilibrium" behavior, whose average play matches the Nash equilibrium, and "strict best response", who best respond to the empirical frequency of her opponent's play. These correspond to the sophisticated and fictitious players analyzed here. Shachat and Swarthout (2004) conduct experiments of a zero-sum game where a human is paired against a computer, which plays a fixed mixed strategy that differs from the equilibrium. They show that the players are able to recognize and exploit the nonequilibrium play. Agents differ, though, in their ability to take advantage of the unsophisticated selections. In more general learning models Camerer and Ho (1998) find heterogeneity in their learning parameters, Camerer, Ho, and Chong (2002) identify a set of sophisticated players who anticipate the potential intelligence of their opponents, and Chong, Camerer, and Ho (2006) distinguish sophisticated from adaptive and myopic agents. These results provide evidence that in models, which allow for heterogeneity in the level of sophistication, explain experimental data well.

Matching Pennies is a zero-sum game. Zero-sum games arise in competitive environments such as competition between firms in a market, politicians in a campaign, negotiators at the bargaining table, criminals and law enforcement officials, and animals in the wild. Understanding how one may use her intelligence against a less sophisticated individual in Matching Pennies provides useful insight into how to behave in such common competitive environments.

II. Analysis

Consider the following two-player, 2x2 normal-form game.

Table 1: Matching Pennies

SP \ FP	1	0
1	w,–w	–w, w
0	–w, w	w,–w

In the one-shot game between two standard, rational players (with $w > 0$) the unique mixed strategy equilibrium is for each player to select 1 with probability ½. This results in an expected payoff to each player of 0.

Consider an infinitely-repeated game where the game depicted in Table 1 is the stage game. Time is indexed by $t = 0, 1, \ldots$ and is discounted at the rate $\delta \in (0,1)$. In each stage each player selects an action from {0,1} and denote s_t as the selection by SP (row player) in stage t. If, within a stage, the players select the same action, then SP receives a payoff $w > 0$ and FP (column player) receives a payoff $-w$. Otherwise, SP receives $-w$ and FP receives w.

Let H_t count the number of times in the first t stages in which SP selects $s_t = 1$, or rather,

$$H_t = \sum_{\tau=0}^{t-1} s_\tau \qquad (1)$$

and denote h_t as the proportion of the first $t > 0$ stages that SP selects t, or rather, $h_t = H_t/t$. Assume that FP is a fictitious player who selects a myopic best response to the mixed strategy of SP that plays 1 with probability h_t and 0 with probability $1 - h_t$. Hence, the play of FP can be described by a function that selects a value of 0 or 1 dependent on the state h_t, denoted $\sigma(h_t)$. It follows from Table 1 that in stage t FP selects $\sigma(h_t) = 1$ if $h_t < ½$ and selects $\sigma(h_t) = 0$ if $h_t > ½$. Assume that if $h_t = ½$, then FP selects $\sigma(h_t) = 1$ with probability $\gamma \in [0,1]$. Furthermore, in $t = 0$ FP also randomizes selecting $\sigma(h_0) = 1$ with probability γ.

Let SP be a standard, sophisticated player. Let p^i denote a path of play for SP, rather $p^i = (s^i_0, s^i_1, s^i_2, \ldots)$. p^i generates a path-specific history $h^i = (h^i_1, h^i_2, \ldots)$ where $h^i_t = (1/t)\sum_{\tau=0}^{t-1} s^i_\tau$. Since FP's choice is a function of the state, h_t, denote SP's expected payoff in stage t as $u(s^i_t,\sigma(h^i_t))$. Consequently, p^i generates a discounted, expected game payoff of

$$U(p^i) = \sum_{t=0}^{\infty} \delta^t u(s_t^i,\sigma(h^i_t)). \qquad (2)$$

SP's objective is to select p^i to maximize $U(p^i)$. At times it will be convenient to consider the continuation payoff of a path from a particular stage. Let $V(h_t)$ denote the continuation payoff to SP if the state is h_t, or rather, for a path p^i in a stage t $V(h^i_t)= \sum_{\tau=t}^{\infty} \delta^{\tau-t}u(s^i_\tau,\sigma(h^i_\tau))$.

Denote P^* as the set of optimal paths, or rather, for $p^* \in P^*$ $U(p^*) \geq U(p^i)$ $\forall p^i$. A path can be classified based on the number of stages in which $h_t = ½$. Let P_n denote the set of paths in which exactly n stages have $h_t = ½$. Hence, P_0 is the set of paths in which h_t never equals ½ and P_∞ is the set of paths that cycle back to the threshold indefinitely.

Additionally, denote $p^e \in P_\infty$ as the path where $s^e_t = 1$ if $h^e_t < ½$, $s^e_t = 0$ if $h^e_t > ½$, and when $\gamma \geq ½$ $s^e_t = 1$ if $h^e_t = ½$ while when $\gamma < ½$ $s^e_t = 0$ if $h^e_t = ½$. Thus, if $\gamma \geq ½$, then $p^e = (1, 0, 1, 0, \ldots)$ and if $\gamma < ½$, then $p^e = (0, 1, 0, 1, \ldots)$. Regardless of γ, $h^e_t = ½$ when t is even and $h^e_t \neq ½$ when t is odd. One may think of this path as a "reversion to the mean" plan where SP looks at her own past play and makes her next selection using an objective of bringing her frequency of play closer to the equilibrium of the one-shot game. As will be shown, one may also think of this as a "one step ahead" plan. Since FP best responds to SP's past play, FP selects the action that does best against the action of SP played most often. Since this is a zero-sum game if SP is able to anticipate this, then selecting the action played less often is a successful plan.

Thus, if SP adopts p^e, then she is looking one step ahead. The following section illustrates that p^e is, in fact, within P^* and, consequently, is the optimal plan for SP to follow.

III. Can You Beat A Fictitious Player?

First, for any finite number of times that the path cycles back to the threshold ½, there is another path that cycles back to the threshold another time that generates a strictly greater game payoff.

Lemma 1: If $p^i \in P_n$ for some $n \neq \infty$, then $p^i \notin P^*$.

Proof. Fix some $n \neq \infty$ and consider a path $p^i \in P^*$. Denote m as the stage where $h^i_m = ½$ and $h^i_t \neq ½$ $\forall t > m$. Consider a path p^j where $s^j_t = s^i_t$ $\forall t \leq m$, $s^i_{m+1} \neq s^i_{m+1}$, and $s^j_t = s^i_{t-1}$ $t > m+1$. Since $V(h^j_m) = V(h^i_m)$, it is required to show that $V(h^j_{m+1}) > V(h^i_{m+1})$. First, by construction $V(h^j_{m+2}) = V(h^i_m)$. It follows that $V(h^j_{m+1}) = w + \delta V(h^i_m)$ and $V(h^i_m) = x + \delta V(h^i_{m+1})$ where x is the payoff received from p^i in t = m. Combining, $V(h^j_{m+1}) = w + \delta x + \delta^2 V(h^j_{m+1})$. Suppose $V(h^i_{m+1}) \geq 0$ (if not then it is straightforward to construct a path that generates a greater game payoff. It suffices to show, then, that $w + \delta x > 0$. Since x is a stage payoff, $x \in [-w,w]$. Since $w > 0$ and $\delta \in (0,1)$, $w + \delta x > 0$. Thus, p^j generates a greater game payoff and $p^i \notin P^*$. This holds for any $p^i \in P_n$ and for any $n \neq \infty$. QED

Consequently, attention can be focused on paths in P_∞. Lemma 2 shows that p^e is an optimal path of play.

Lemma 2: $U(p^e) \geq U(p^i)$ $\forall p^i \in P_\infty$.

Proof. Consider, first, $p^m \in P_\infty$ where $s^m_t = 1$ if $h^e_t < ½$, $s^m_t = 0$ if $h^e_t > ½$, and when $\gamma \geq ½$ $s^m_t = 0$ if $h^m_t = ½$ while when $\gamma < ½$ $s^m_t = 1$ if $h^m_t = ½$ (so that, like p^e, $h^m_t = ½$ in every even t, but the opposite selection is made in the even numbered stages). Thus, p^m generates w in every odd t (just as p^e does), but in every even t receives, if $\gamma < ½$, $\gamma w + (1-\gamma)(-w) < 0$ and, if $\gamma \geq ½$, $\gamma(-w) + (1-\gamma)w \leq 0$. Thus, $U(p^e) \geq U(p^m)$ (and this holds with equality if and only if $\gamma = ½$). Similarly, any other path where $h_t = ½$ for even t that is not p^e generates a lower game payoff. Therefore, consider a path $p^n \in P_\infty$ that does not have $h^n_t = ½$ for every even t. Since in t = 0 $\sigma(h_0) = 1$ with probability γ, the proof of Lemma 1 can be applied. As shown in Lemma 1, in any even-numbered stage a deviation selecting the action that brings h_t closer to ½ is a profitable one. Hence, $U(p^e) \geq U(p^n)$ and, since this covers all possible paths in P_∞, $U(p^e) \geq U(p^i)$ $\forall p^i \in P_\infty$. QED

Now turn to the main result.

Proposition 1: $p^e \in P^*$.

Proof. From Lemma 1, since $p^* \notin P_n$ $\forall n \neq \infty$, then $p^* \in P_\infty$ $\forall p^* \in P^*$. From Lemma 2 this implies $U(p^e) \geq U(p^*)$. By definition of P^* $U(p^e) \leq U(p^*)$ $\forall p^* \in P^*$. Consequently, $U(p^e) = U(p^*)$ and $p^e \in P^*$. QED

The uniqueness of p^* can be determined.

Corollary 1: If $\gamma \neq ½$, then P^* is a singleton.

Proof. From the proof of Lemma 2 if $\gamma \neq ½$, then $U(p^e) > U(p^i)$ $\forall p^i \in P_\infty$, $p^i \neq p^e$. From Lemma 1 if $\exists$ $p^* \in P^* \ni p^* \neq p^e$, then $p^* \in P_\infty$. Thus, $U(p^e) > U(p^*)$, but by definition of P^* $U(p^*) \geq U(p^e)$. Due to this contradiction, if $\gamma \neq ½$ and $p^* \in P_\infty$, then $p^e = p^*$ and, as a consequence, p^e is the only element in P^*.

QED

In odd-numbered stages $h^e_t \neq ½$. If $h^e_t > ½$, then $s^e_t = 0$ and $\sigma(h^e_t) = 0$. Hence, $u(0,0) = w$. If $h^e_t < ½$, then $s^e_t = 1$ and $\sigma(h^e_t) = 1$ so that $u(1,1) = w$. Consequently, in every odd-numbered stage SP receives w. Additionally, $h^e_t = ½$ for all even stages. Thus, $u(s^e_t,\sigma(h^e_t)) = \gamma w + (1-\gamma)(-w)$ if $\gamma \geq ½$ and $\gamma(-w) + (1-\gamma)w$ if $\gamma < ½$. Hence, $u(s^e_t,\sigma(h^e_t)) = |1 - 2\gamma|w \equiv z \in [0,w]$ for each even-valued t. As a consequence, $U(p^e) = z + \delta w + \delta^2 z + \delta^3 w + ...$, or rather,

$$U(p^e) = (z+\delta w)/(1-\delta^2) > 0. \quad (3)$$

Thus, SP does strictly better than the game payoff she would expect to receive playing against another sophisticated player, which is $0/(1-\delta) = 0$. Additionally, if FP's randomization is the same as the mixed strategy equilibrium, $\gamma = ½$, then $z = 0$ and SP's payoff is smallest, but still greater than zero. If FP's indifference-breaking behavior is a pure action, $\gamma = 0$ or 1, then $z = w$ and SP wins in every stage obtaining her maximax payoff.

Finally, in adaptive models it is common to consider the long run outcome. If one considers convergence in beliefs, then $h^e_t = (t+I)/2t$ where $I = 1$ if t is odd and $\gamma \geq ½$, $I = 0$ if t is even, and $I = -1$ if t is odd and $\gamma < ½$. The sequence $\{(t+I)/2t\}$ converges to ½ as $t \to \infty$. Hence, play converges in beliefs to the mixed strategy Nash equilibrium even though the outcome diverged from it.

IV. How Sophisticated Must You Be?

As a consequence, the intelligence of the sophisticated player works to her benefit. One may ask, then, just how intelligent must she be?

The optimal path alternates between the two actions. As discussed previously, one way to interpret this is to suppose that SP looks "one step ahead". SP may be able to recognize that at the given state the opponent is going to play a myopic best response to it. Knowing this she takes advantage of the myopia. While FP considers the selections made in stages 0, 1, 2, ..., t-1, if SP additionally considers stage t, then the greatest payoff possible can be achieved. She does not need to act with any more foresight than this.

Another interpretation is that SP plays a "reversion to the mean" plan. She acknowledges that the equilibrium in the game (with rational players) is to play each action with probability ½. If she has selected an action in less than one-half of the stages, then it is chosen. If a player adopts such a behavior, then she, in fact, achieves the same expected game payoff as if she is a sophisticated player. Mookherjee and Sopher (1994) analyze experiments of Matching Pennies and, while the empirical frequency of play converged to the mixed strategy Nash equilibrium, they find evidence of play that is not i.i.d. Rather, subjects exhibited negatively autocorrelated strategies. Walker and Wooders (2001) consider data from tennis serves at Wimbledon and, again, find the empirical frequency of play near the mixed strategy equilibrium, but with negatively autocorrelated strategies. Thus, if a player attempts to keep the frequency of play close to the mixed strategy distribution and switches too often, then she may be able to do very well against a fictitious player.

V. Conclusion

So, can you beat a fictitious player at Matching Pennies? Yes, your expected game payoff would be strictly greater than that received playing against another sophisticated player. Just how much greater depends on the randomizing behavior of the fictitious player if it is indifferent. The farther this is away from the mixed strategy Nash equilibrium the greater is your game payoff and, in fact, if a pure strategy is selected to break the indifference, then you get your maximax payoff. The optimal path switches the action selected in every stage. If the number of times an action has been selected is less than one-half of the number of possible stages, then it is chosen and wins. In the stages where each has been played an equal number of times the action chosen is the one that is a myopic best response to the randomization of the fictitious player. Hence, a player need only select a rather simple plan to beat the fictitious player. This illustrates the importance of the assumption of symmetry of intelligence. Additional consideration for different adaptive behaviors and other types of games is needed.

VI. References

Boylan, R., and M. El-Gamal (1992) "Fictitious Play: A Statistical Study of Multiple Economic Experiments" *Games and Economic Behavior* 5, 205-222.

Brown, G. (1951) "Iterative Solution of Games by Fictitious Play" in *Activity Analysis of Production and Allocation: Proceedings of a Conference* by T. C. Koopmans, Ed., Wiley: New York.

Cahn, A. (2004) "General Procedures Leading to Correlated Equilibrium" *International Journal of Game Theory* 33, 21-40.

Camerer, C., and T.-H. Ho (1998) "Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation" *Journal of Mathematical Psychology* 42, 305-326.

Camerer, C., T.-H. Ho, and J.-K. Chong (2002) "Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games" *Games and Economic Behavior* 104, 137-188.

Chong, J.-K., C. Camerer, and T.-H. Ho (2006) "A Learning-Based Model of Repeated Games with Incomplete Information" *Games and Economic Behavior* 55, 340-371.

Ellison, G. (1997) "Learning from Personal Experience: One Rational Guy and the Justification of Myopia" *Games and Economic Behavior* 19, 180-210.

Fudenberg, D., and D. Levine (1995) "Consistency and Cautious Fictitious Play" *Journal of Economic Dynamics and Control* 19, 1065-1089.

Fudenberg, D., and D. Levine (1998) *Theory of Learning in Games*, MIT Press: Cambridge.

Krishna, V., and J. T. Sjöström (1997) "Learning in Games: Fictitious Play Dynamics" in *Cooperation: Game-Theoretic Approaches* by S. Hart and A. Mas-Colell, Eds., Springer-Verlag: New York.

Miyasawa, K. (1961) "On the Convergence of the Learning Process in a 2x2 Non-Zero-Sum Two-Person Game" Princeton University, Memo No. 33.

Mookherjee, D., and B. Sopher (1994) "Learning Behavior in an Experimental Matching Pennies Game" *Games and Economic Behavior* 7, 62-91.

Ochs, J. (1995) "Games with Unique, Mixed Strategy Equilibria: An Experimental Study" *Games and Economic Behavior* 10, 202-217.

Robinson, J. (1951) "An Iterative Method of Solving a Game" *Annals of Mathematics* 54, 296-301.

Shachat, J., and J. T. Swarthout (2004) "Do We Detect and Exploit Mixed Strategy Play By Opponents?" *Mathematical Methods of Operations Research* 59, 359-373.

Vriend, N. (1997) "Will Reasoning Improve Learning?" *Economic Letters* 55, 9-18.

Walker, M., and J. Wooders (2001) "Minimax play at Wimbledon" *American Economic Review* 91, 1521-1538.