ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 69, No. 1, March, pp. 31–49, 1997
ARTICLE NO. OB962670

Randomization and Adaptive Learning in a Simplified Poker Game

AMNON RAPOPORT
University of Arizona

IDO EREV
Technion–Israel Institute of Technology, Israel

AND

ELIZABETH V. ABRAHAM AND DAVID E. OLSON
University of Arizona

Thirty pairs of subjects participated in three different two-person zerosum Poker games in extensive form with imperfect and asymmetric information. The results provide no support for the (unique) mixed-strategy equilibrium solution for risk-neutral players on either the individual or the aggregate level. Compared to this solution, the informed players do not bluff as often as they should, and the uninformed players call too often. Comparison of the present study with previous studies of diametrically opposed conflicts shows consistent differences between the extensive form of two-person zerosum games and its strategically equivalent normal form. An adaptive learning model proposed by Roth and Erev (1995) succeeds in tracking the general trends in the probabilities of bluffing and calling over time and in accounting for the effects of size of stake, when it assumes that learning takes place only on stage games when the players are called upon to act.

One of the earliest and most thorough investigations of Poker appears in the classical treatise on game theory—*Games and Economic Behavior* by von Neumann and Morgenstern (1947). A large section of Chapter 4 is devoted to the formal analysis of "bluffing" in several simplified variants of a two-person Poker game with either symmetric or asymmetric information. Indeed, the general considerations concerning Poker and the mathematical discussions of the variants of the game analyzed in this chapter were carried out by von Neumann as early as 1926. Recognizing that "bluffing" in Poker "is unquestionably practiced by all experienced players," von Neumann and Morgenstern identified two motives for bluffing. "The first is the desire to give a (false) impression of strength in (real) weakness; the second is the desire to give a (false) impression of weakness in (real) strength" (von Neumann & Morgenstern, 1947, p. 189).

Solutions to these variants of the simplified Poker game as well as a large class of both zerosum and nonzerosum games were unified by the concept of mixed strategy—a probability distribution over the player's set of actions. The importance of mixed strategies to the theory of games and its applications to the social and behavioral sciences stems from the fact that for many interactive decision situations modeled as noncooperative games, both zerosum and nonzerosum, there are no Nash equilibria in pure strategies. The major problems associated with mixed strategies as solutions to interactive decision problems are that they are often counterintuitive (Rubinstein, 1991), unstable (Stahl, 1988), and have limited appeal in many practical applications (Radner & Rosenthal, 1982).

These problems have given rise to two major issues, one theoretical and the other empirical. The theoretical issue concerns the *interpretation* of mixed strategies. A mixed strategy seems to entail a deliberate and conscious attempt by the player to introduce randomness into her behavior, thereby committing herself to a random device with prespecified probabilities (e.g., a biased coin) which probabilistically selects members of her set of actions (Osborne & Rubinstein, 1994). Two arguments have been raised against this interpretation. The first argument is that, in actuality, decision makers do

The first two authors acknowledge support for this project from the Israel Science Foundation (Grant 604-94-1). Address correspondence and reprint requests to Amnon Rapoport, Department of Management and Policy, University of Arizona, McClelland Hall 405, Tucson, AZ 85721. Fax: (520)621-4171. E-mail: amnon@aruba.ccit.arizona.edu.

0749-5978/97 $25.00

not rely on random devices to determine their action. The second argument questions the need to randomize in the first place. In equilibrium, each of the players is necessarily indifferent among the pure strategies to which the recommended mixed strategy attaches positive probabilities. Therefore, there is no particular reason why the player ought to play the equilibrium strategy (Binmore, 1992).

Attempts to counter these arguments have proposed different interpretations of mixed strategy (e.g., Osborne & Rubinstein, 1994), which do not require the players to make deliberate effort to choose their actions with the probability weights prescribed by their equilibrium mixed strategies. Noted among them are the "purification" rationale of Harsanyi (1973) and the "equilibrium in beliefs" interpretation (see, e.g., Binmore, 1992). Although other interpretations have been proposed (see, e.g., Osborne & Rubinstein, 1994, for a detailed discussion), there seems to be no consensus about the classes of interactive decision making situations to which any one of them applies. Moreover, there seems to be an agreement that at least in certain interactive situations randomization over pure strategies is a *deliberate* action rather than the outcome of variations in payoffs or beliefs. Osborne and Rubinstein write, "For example, players randomly 'bluff' in poker, governments randomly audit taxpayers, and some stores randomly offer discounts" (1994, p. 37). And, in commenting on the purification idea, Binmore writes "This is an attractive idea but it should not be allowed to displace other interpretations altogether. In particular, Poker players will often be wise to randomize very actively" (1992, p. 520).

The second major issue is the *empirical validity* of mixed strategies. This issue is important because so many models in the social and behavioral sciences are based on Nash equilibria. Several questions suggest themselves: Do people randomize their actions? And, if they do, do they randomize with the probability weights prescribed by their mixed equilibrium strategies? Do people exploit deviations from equilibrium by their opponents? With repetition of the stage game, do the strategy choices of the players converge to the equilibrium mixed-strategy profile? Answers to these and related questions depend on a variety of experimental design issues including characteristics of the game (e.g., number of players, form of presentation, number of pure strategies to each player, zerosum vs. nonzerosum games, one-shot vs iterated games), risk attitude, control over players' motivation, and procedures used to induce mixed strategies. It is no wonder that answers to the questions posed above cannot be settled by a single experiment but, rather, can only emerge slowly from a variety of experiments that differ from one another in their design characteristics and population of subjects.

Previous studies, to be briefly reviewed below, have investigated randomization in two-person zerosum games presented in strategic (normal) form. In contrast, most conflicts evolve in real time. The major purpose of the present study is to extend these studies to games calling for sequential rather than simultaneous play. It is, of course, rather straightforward to transform two-person zerosum games from extensive to strategic (normal) form by using the game theoretical concept of strategy. However, as noted by Kreps (1990), Rubinstein (1991), Greenberg (1994), and many others, the notion of equilibrium mixed strategies is particularly problematic in extensive form games with imperfect information. Moreover, recent work by Rapoport (in press), Rapoport and Fuller (1995), and Güth, Huck, and Rapoport (in press), which shows systematic effects of priority in time (which is not associated with priority in information) in games with imperfect information, suggests that the extensive form of a game, which calls for sequential play, and its strategically equivalent normal form, which calls for simultaneous play, may give rise to different behavioral regularities. This may be particularly the case in extensive form games, like our simplified Poker game, which allow for certain signals (e.g., bluffs) whose interpretation may be susceptible to social norms or moral considerations.

Our purpose is to study two-person games with diametrically opposed interests evolving in real time which call for randomization of pure strategies by each player. We propose first to test the game theoretical hypothesis that players randomize their strategies with the probability weights prescribed by their mixed equilibrium strategies. The second hypothesis, which is also implied by game theory, is that behavior in extensive form games does not differ from behavior in their strategically equivalent games presented in normal form. Thus, even if the first hypothesis is rejected, deviations from equilibrium under sequential play should not differ from deviations from equilibrium under simultaneous play. Finally, we propose to test the hypothesis that the adaptive learning model proposed by Roth and Erev (1995), which accounts rather well for the main regularities in a variety of experimental studies of 2×2 zerosum games in normal form (Erev & Roth, 1995), can also account for the dynamics of play in our study.

PREVIOUS STUDIES OF TWO-PERSON GAMES IN STRATEGIC FORM

There are at least three major reasons why two-person zerosum games have been a major tool for studying randomization in strategic environments. First, if there exists a mixed-strategy equilibrium in two-person zerosum games, it is unique. Therefore, problems of multiplicity of equilibria do not arise. Second, if the game is iterated in time with fixed pairs, there should be no reason—so the theory says—to worry about reputation effects or tacit collusion. Third, as discussed below, the issue of utility measurement can be bypassed by constructing payoff matrices with only two payoffs for each player.

Experiments that have subjects repeatedly play two-person zerosum games (with no pure strategy equilibrium) in strategic form were already conducted in the 1960s and 1970s (e.g., Fox, 1972; Lieberman, 1960, 1962; Messick, 1967; Morin, 1960, Suppes & Atkinson, 1960). These experiments have yielded mixed results with the bulk of evidence disconfirming the minmax solution as a descriptive model. O'Neill (1987) criticized these experiments, which have all assumed that utility depends only on the player's own payoff and is a linear function of that payoff. To resolve these methodological problems, O'Neill proposed to construct payoff matrices with only two payoffs for each player. The invariance over the two utility functions, one for each player, occurs because there are only two outcomes, a win and a loss, and the minmax solution is known to be invariant under a positive linear transformation of each player's utilities. The same approach has been used in subsequent experiments on mixed strategies in repeated two-person zerosum games conducted by Budescu and Rapoport (1994), Mookherjee and Sopher (1994), Shachat (1995), Rapoport and Budescu (1992), and Rapoport and Boebel (1992).

To summarize the major findings of these experiments, we shall refer to the relative frequencies of the pure strategies summed across all the iterations of the stage game as the *observed* mixed strategy, and to the theoretical probabilities as the *predicted* mixed strategy. We shall also differentiate between mixed strategies computed on the individual or aggregate level. With this terminology, the major findings of these experiments can be summarized as follows:

1. Players mix their pure strategies as expected, but the observed individual mixed strategies do not conform exactly to the predicted mixed strategy.

2. Individual differences between the observed mixed strategies are substantial. Because of the payoff structure of the games, these individual differences cannot be accounted for by postulating different utility functions.

3. The observed aggregate mixed strategies for risk-neutral players approximate the predicted mixed strategy quite well. The approximation gets worse as the size of the payoff matrix grows larger.

4. The null hypothesis that the individual sequence of choices elicited from the player in the iterated game is consistent with the pattern that would be generated by a common, stationary, multinomial model implied by the minmax theorem is rejected for most of the players (Brown & Rosenthal, 1990; Rapoport & Boebel, 1992; Shachat, 1995). In particular, there is evidence for a small but persistent serial dependence in the individual choices which results in *overalternation*.

5. Players change the mixing of their strategies across iterations (see, in particular, Binmore, Swierzbinski, & Proulx, 1996). This adaptive learning explains in part finding 4. above and may account in part for the differences between dyads.

6. When compared with other static models, the minmax hypothesis provides a better account of the results (see, e.g., Binmore *et al.*, 1996; Rapoport & Boebel, 1992).

THE SIMPLIFIED POKER GAME

The simplified Poker game presented below is a two-person game in extensive form with diametrically opposed preferences. It can be described in terms of five steps, which paraphrase the instructions given to the subjects (see below).

Step 1. Each of two players, named A and B, puts a single poker chip in a bowl placed on a table between the two players.

Step 2. Player A draws a card from a well-shuffled deck with 75 cards marked H (for "High") and 75 marked L (for "Low"). Player A observes the letter on the card, records it, and then places the card face down on a desk between the two players. Player B is informed of the number of H and L cards in the deck, but not of the actual outcome of the drawing.

Step 3. Having observed the letter on the card, Player A makes one of two decisions called "Drop" (D) and "Raise" (R). If he drops, Player B takes the two chips from the bowl and the game is over. If he raises, Player A adds x more chips to the bowl.

Step 4. Having observed Player A raising the stakes, Player B can either "Fold" (F) or "Call" (C). If she folds, Player A takes the $x + 2$ chips from the bowl and the game is over. If she calls, Player B adds x more chips to the bowl.

Step 5. If Player A raises and then Player B calls, the card is turned face up. If it is marked H, Player A takes all the chips in the bowl (a total of $2x + 2$), and the game is over. If it is marked L, the chips in the bowl are claimed by Player B.

Because Player A observes the letter on the card and Player B does not, henceforth we shall occasionally refer

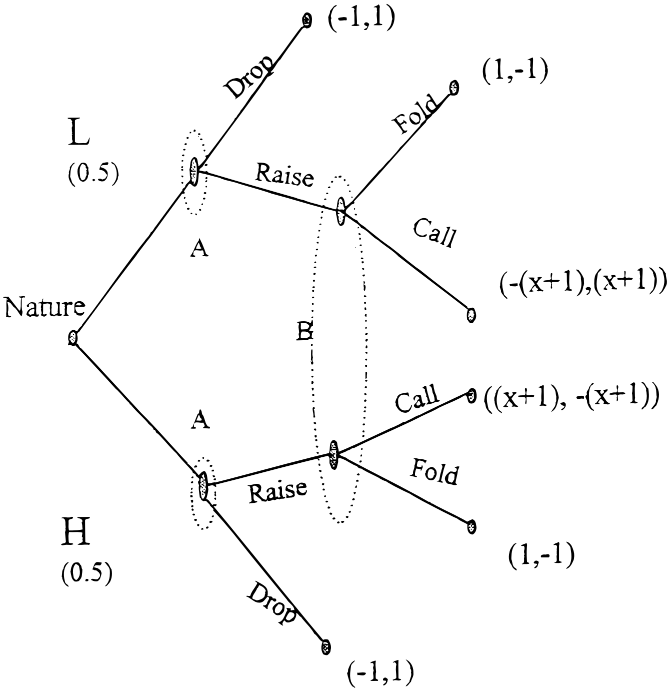

FIG. 1. The simplified poker game in extensive form.

to Players A and B as the "informed" and "uninformed" players.

Figure 1 portrays the Poker game in extensive form. The (equal) probabilities of the two types of cards H and L (in parentheses) are common knowledge. The game is zerosum in terms of the players' actual payoffs, which assume the values 1, -1, $(x + 1)$, and $-(x + 1)$, but not necessarily in terms of the players' utilities. Player A has four pure strategies and Player B just two.

The strategic form of the game is shown in Table 1; the entries in each cell are the expected utilities of Players A and B, respectively. Assuming that $-u(-x) = u(x)$ and $u(0) = 0$, it is easy to verify that strategy (R,R) ("always raise") of Player A strictly dominates strategy (D,D) ("always drop") as well as strategy (R,D) ("raise on L, drop on H"). Elimination of these two strictly dominated strategies yields a 2×2 irreducible payoff matrix, which has no equilibrium in pure strategies. Solving for the (unique) mixed strategy equilibrium, we obtain the following solution for Player A:

$$\text{prob(R,R)} \equiv p = [u(x + 1) - u(1)]/[u(x + 1) + u(1)] \tag{1}$$

and

$$\text{prob(D,R)} \equiv 1 - p = 2u(1)/[u(x + 1) + u(1)].$$

The equilibrium solution for Player B is:

$$\text{prob(C)} \equiv q = 2u(1)/[u(x + 1) + u(1)] \tag{2}$$

and

$$\text{prob(F)} \equiv 1 - q = [u(x + 1) - u(1)]/[u(x + 1) + u(1)].$$

In equilibrium, Player A should always raise on H, but only raise on L ("bluff") with probability p. Having observed Player A raising the stakes, Player B should call with probability q and fold with probability $1 - q$. Note that if the utility functions of the two players are identical, then $p = 1 - q$ for each value of x. Note, too, that the probability of raising on L, denoted by p, increases in x, whereas the probability of calling, denoted by q, decreases in x.

Under the assumption that each player's utilities are linear in money, Eqs. (1) and (2) reduce to

$$p = 1 - q = x/(x + 2), \tag{3}$$

and

$$q = 1 - p = 2/(x + 2).$$

The value of the game (for Player A) is then $x/(x + 2)$.

METHOD

Subjects

Sixty subjects, mostly undergraduate and graduate students at the University of Arizona, participated in the experiment. Of these 60 subjects, 22 were females and 38 were males. The subjects volunteered to take

TABLE 1

The Simplified Poker Game in Strategic Form

Player A	Player B: F (Fold)		C (Call)	
(D,D)	u(−1),	u(1)	u(−1),	u(1)
(D,R)	[u(−1) + u(1)]/2,	[u(−1) + u(1)]/2	[u(−1) + u(x + 1)]/2,	[u(1) + u(−(x + 1))]/2
(R,D)	[u(1) + u(−1)]/2,	[u(−1) + u(1)]/2	[u(−(x + 1)) + u(−1)]/2,	[u(x + 1) + u(1)]/2
(R,R)	u(1),	u(−1)	[u(−(x + 1)) + u(x + 1)]/2,	[u(x + 1) + u(−(x + 1))]/2

part in a decision making experiment with payoff contingent on performance, which was advertised on bulletin boards across campus. The subjects participated in three experimental conditions in a between-subjects design with 20 subjects (10 dyads) in each condition (see below). They were instructed to expect earning at least $10 and possibly much more contingent on their performance. No previous experience in playing Poker was required. The experiment lasted approximately 100 min.

Procedure

The experiment was conducted by two experimenters in two separate rooms. To enhance the players' trust in the validity of the procedure (by ensuring that subjects actually observe the randomization process), and allow comparison of our findings with the results of previous studies of mixed strategies in two-person zerosum games in strategic form, most of them using fixed pairs (but see Shachat, 1995, for a different design), interaction within dyad was face-to-face. On each session, four subjects divided into two dyads were run simultaneously. Within each dyad, the player's role (A or B) was determined by the toss of a fair coin.

The two dyad members were seated on two opposing sides of a table. Each was handed a set of written instructions and a stack of poker chips. An empty bowl and a stack of 150 well-shuffled cards, 75 marked H and 75 marked L, were placed in the middle of the table between the two players.

After the subjects had completed reading the instructions, the experimenter answered their questions, if any. Then the stage game commenced with each player taking one of the chips from his stack and putting it in the empty bowl. Player A then drew a card from the top of the deck, observed its label (H or L), and placed it on the table face down. The stage game then proceeded according to the five steps of the Poker game described above. The card drawn by Player A, as well as the actions taken by the players (D or R by Player A and F or C by Player B), were recorded by the experimenter. Cards were replaced and reshuffled.

The two stacks of chips were visible to both players during the entire duration of the experiment. On each stage game, one of the two players transferred to the other either 1 or $x + 1$ chips, depending on the outcome of the game. The experiment continued in this manner, with chips switching hands on each period, for a prespecified (and commonly known) number of periods or until one of the players ran out of chips, whichever occurred first. This concluded Phase 1 of the experiment.

At the end of Phase 1, each dyad member was asked to tally his chips and report the number to the experimenter. The players were then informed that, to be fair, each would have the opportunity to repeat the experiment in the opposite role. However, the opponents were switched with the members of the other dyad run by the second experimenter, so that each player faced a new opponent in a new role. Other than switching partners and roles, Phase 2 was identical to Phase 1.

TABLE 2

Parameters for the Three Experimental Conditions

Experimental condition	No. of periods	Stack size of A	Stack size of B	x	Value of game	Mean payoff
C2	72	24	96	2	1/2	$17.00
C4	120	30	190	4	2/3	$22.00
C6	120	30	210	6	3/4	$24.00

Upon completion of both phases, the number of chips gained by each subject in both phases was tallied. The chips were then converted into money according to a conversion rate that had been told to the subjects at the beginning of the experiment. The subjects were paid separately and dismissed.

Design

The three experimental conditions, labeled C2, C4, and C6, varied from one another in the value of x; x = 2, 4, and 6 for Conditions C2, C4, and C6, respectively. The stacks of poker chips given to the dyad members differed in size. Their sizes were determined so that, if both players adhered to the equilibrium strategy (assuming risk neutrality) and no player ran out of chips, both would end up with approximately the same number of chips.

Table 2 presents the number of periods in each phase, stack sizes for both players, value of x, value of the stage game, and mean payoff by condition. For example, each subject played 120 iterations of Condition C6 as Player A and 120 iterations (against a different opponent) as Player B for a total of 240 periods. The stack sizes in this condition were 30 and 210 poker chips for Players A and B, respectively. Note that the value of the stage game in this condition is 3/4 chip. Therefore, if both dyad members play their equilibrium strategies and neither player runs out of chips, Player B should transfer to player A on the average 90 chips, and each should end up the phase with approximately 120 chips.[1]

[1] There is no doubt that the unequal endowments signaled to the subjects information about their relative strengths and the advantage that the informed player had over the uninformed player. A subsequent study in which the initial stakes are equalized is clearly warranted.

Our decision to restrict the initial number of chips given to the players was made to enhance competition. A similar procedure was used by Rapoport and Boebel (1992). Bankruptcy occurred very infrequently. In Condition C2 all the 20 dyads completed 72 periods. In Condition C4 one of the 20 dyads completed only 116 periods and the remaining 19 dyads each completed all 120 periods. Finally, in Condition C6 one dyad completed 108 periods, another dyad completed 109 periods, and the remaining 18 dyads each completed all the 120 periods.[2]

RESULTS

To check whether the sequence of cards drawn by Player A was, in fact, a realization of a Bernoulli process with equally likely alternatives, we calculated for each dyad separately the proportion of H cards. Statistical tests (using the normal approximation to the Binomial distribution with $p = .05$) failed to reject the null hypothesis $p(\mathrm{H}) = p(\mathrm{L}) = .5$ in 58 of the 60 cases (20 dyads in each of the three conditions). Therefore, we have no evidence that the cards drawn and observed by Player A were not well shuffled.

The analyses of the remaining results are presented in two sections. Focusing on static aspects of the data, the first section reports summary statistics on both the individual and aggregate levels and then uses them to test several qualitative and quantitative implications of the equilibrium solution. In the second section we attempt to account for the dynamic aspects of the data with a simple adaptive learning model recently proposed and tested by Roth and Erev (1995).

Qualitative and Quantitative Implications

Dropping on H is a dominated strategy for Player A. With very few exceptions that may be attributed to error, type A players always raised on H. The number of players of type A (out of 20) who always raised on H was 15, 14, and 16 in Conditions C2, C4, and C6, respectively. The (estimated) mean probability of raising on H, computed across the 20 type A players in each condition, was .992, .983, and .995 in Conditions C2, C4, and C6, respectively. The median in each condition was 1. As predicted, with very few exceptions dominated strategies ((D,D) and (R,D)) were never chosen by Player A.

Continuing the analysis, we estimated for each type A player the probability of raising on L, which we previously (Eq. 1) denoted by p. This computation was based only on the number of periods on which type A players actually observed a low card (approximately 50% of all periods). In a similar way, we estimated the probability of calling, which we previously denoted (Eq. 2) by q, for type B players. Because type A players almost always raised on H and in approximately 50% of the periods raised on L, the probability q is based on approximately 75% of all periods. Table 3 presents the means and standard deviations of the estimated p and q values by condition. It also displays the means and standard deviations of the sum of these two probabilities, namely, $p + q$.

For both types of players and for all three conditions, considerable individual differences in the values of p and q were observed. The individual values of p ranged between .24 to .68 in Condition C2, .07 to .69 in Condition C4, and 0 to .81 in Condition C6. Correspondingly, the values of q ranged between .45 to .90 in Condition C2, .24 to .83 in Condition C4, and .06 to .76 in Condition C6. To test for phase effects, the 10 values of p in Phase 1 were compared to the 10 values of p in Phase 2 for each condition separately by t test. The results were mixed. There was no significant phase effect for Conditions C2 and C6, but the phase effect in Condition C4 was significant ($p < .05$). On the average, type A players in Condition C4 raised on L more often in Phase 2 than in Phase 1. We repeated the same analysis comparing the 10 q values in Phase 1 with the 10 q values in Phase 2. None of the three t tests yielded significant differences between the mean q values ($p = .05$).

In equilibrium, the value of p should increase in x, whereas the value of q should decrease. Table 3 shows no evidence that, on the average, the probability of raising on L increased as the size of the stake grew larger. On the average, in all three conditions type A players were as likely to raise on L as to drop. In contrast, the probability of calling by type B players decreased from .70 in Condition C2, through .55 in Condition C4, to .50 in Condition C6.

With the actual payoffs used in our experiment, we have no way of bypassing the problem of utility assessment. However, we can derive predictions which are

TABLE 3

Means and Standard Deviations of Probabilities of Raising (R|L) and Calling (C|R) by Condition

	Condition C2		Condition C4		Condition C6	
Prob.	Mean	*SD*	Mean	*SD*	Mean	*SD*
p	0.47	0.13	0.50	0.19	0.47	0.24
q	0.70	0.11	0.55	0.19	0.50	0.21
$p + q$	1.17	0.32	1.05	0.37	0.97	0.42

[2] Strictly speaking, if one of the players can get bankrupt, the stage game is no longer zerosum as the outcome includes the player's stage game payoff as well as the probability of continuation. However, in our experiment this constraint was hardly binding as only 3 of the 60 games were not completed.

utility invariant for *pairs* of players. Equations (1) and (2) imply that, under the condition that both dyad members share the same attitude toward risk and as long as $u(x) = -u(-x)$, $p + q = 1$ regardless of the shape of the (common) utility function. The estimated sum of probabilities $p + q$ was computed for each dyad separately. The means and standard deviations are presented in Table 3. We then tested the null hypothesis $p + q = 1$ against the alternative two-sided hypothesis for each condition separately. We rejected the null hypothesis in Condition C2 ($t(19) = 2.38$, $p < .01$) but not in Conditions C4 ($t(19) = .56$) and C6 ($t(19) = .27$). These results suggest that, on the average, dyads in Conditions C4 and C6 found (p, q) equilibria which might correspond to different utility functions across dyads.

Utility analysis. We proceeded by attempting to account for the estimated probabilities p and q by making specific assumptions about the players' utility functions. Under the assumption of risk neutrality, Equation (3) implies that p should assume the values 1/2, 2/3, and 3/4 in Conditions C2, C4, and C6, respectively. Whereas the mean p value supports this prediction in Condition C2, it disconfirms it in the other two conditions. Under the same assumption, q should assume the values 1/2, 1/3, and 1/4 in Conditions C2, C4, and C6, respectively. Table 3 shows that, compared to the equilibrium solution (Eq. 3), type B players called much too often. On the aggregate level, these results could be accounted for by assuming that Player A responds to Player B's behavior. If, in comparison to the equilibrium solution, Player B calls too often, then Player A should decrease his probability of raising on L. These results exclude the other possibility; if Player A raises less often than predicted, then Player B should decrease the probability of calling. The actual payoffs are in agreement with the hypothesis that the informed players respond to the uninformed players' behavior and improve their position. The mean final number of chips owned by Player A at the end of the game in Conditions C2, C4, and C6 was 58.6, 124.4, and 142.0, respectively. In contrast, the corresponding number of chips owned by Player B was 61.4, 95.7, and 98.1. On the average, the informed players were paid more than the uninformed players (and the equilibrium solutions) in Conditions C4 and C6.

To test the equilibrium solution for risk-neutral players on the individual (rather than aggregate) level, we compared the observed[3] and predicted (Eq. 3) p values for each type A player separately (using a t test). The equilibrium solution was rejected on the individual level ($p < .05$) for 4 of the 20 subjects in Condition C2, 14 subjects in Condition C4, and 14 subjects in Condition C6. In a similar vein, we compared the observed and predicted (Eq. 3) q values for each type B player. Out of 20 t tests, the equilibrium solution (assuming risk neutrality) was rejected in 14, 14, and 16 cases in Conditions C2, C4, and C6, respectively. With only a few exceptions, we find no evidence that the equilibrium solution with linear utilities could account for the individual results across all trials.

As an alternative hypothesis to risk-neutrality, we computed and tested the equilibrium solution in mixed strategies assuming the "utility" function

$$u(x) = \begin{cases} x^a, & \text{if } x \geq 0 \\ -c(-x)^a, & \text{if } x < 0 \end{cases}$$

proposed by Prospect Theory (Tversky & Kahneman, 1992). In this function, the parameter "a" represents risk attitude, and the parameter "c" loss aversion (if $c \geq 1$). If both dyad members share the same function, then it can be shown that in equilibrium

$$p = c[(x + 1)^a - 1]/[(x + 1)^a + c]$$

and

$$q = (1 + c)/[c(x + 1)^a + 1].$$

Once p and q are estimated from the data, these two equations can be solved simultaneously for the parameters c and a. The solution is consistent with Prospect Theory if $0 \leq a < 1$ and $c \geq 1$. Given the estimated values of p and q, we computed the parameters c and a for each dyad separately. Results consistent with Prospect Theory, which assumes utility functions concave in the domain of gains, convex in the domain of losses, and steeper in the latter domain, were found for only three, two, and two dyads in Conditions C2, C4, and C6, respectively. We find no evidence that the equilibrium solution with Prospect Theory type of utility accounts for the dyadic results.

Sequential dependencies. If adaptive play *does not* take place in the iterated game—a strong and probably invalid assumption—the equilibrium solution implies that, regardless of his/her attitude toward risk, each player should generate a binary sequence of responses (D and R for Player A and C and F for Player B). "Randomization" in this case means that the sequence produced does not differ significantly from a Bernoulli series. This null hypothesis has several implications, one

[3] Each p value is calculated on the basis of all the trials on which a low value card was observed; there are approximately 36, 60, and 60 trials for Players A in Conditions C2, C4, and C6, respectively.

of which concerns the first-order sequential dependencies. If Player A produces a Bernoulli series, then the conditional probability of raising on L, given that he raised the stakes on the immediately preceding period in which he drew a low card, denoted by $p(R|R)$, should not differ from the conditional probability of raising on L, given that he dropped on the last period that he drew a low card, $p(R|D)$. To test the null hypothesis $p(R|R) = p(R|D)$, we computed for each type A player these two first-order conditional probabilities and subjected them to a sign test (one for each condition). The results show that $p(R|D) > p(R|R)$ for 18 of the 20 type A players in condition C2, 17 of the 20 type A players in Condition C4, and 15 of the 20 type A players in Condition C6. The null hypothesis was, therefore, clearly rejected (at $p < .05$) in each condition. As in all previous studies of mixed strategies in two-person zerosum games in strategic form, most of our type A players produced non-Bernoulli type series exhibiting a strong overalternation bias.

We repeated the same analysis on the second-order conditional probabilities, and tested separately for each condition two null hypotheses: $p(R|R, R) = p(R|R, D)$ and $p(R|D, R) = p(R|D, D)$, where $p(R|R, R)$ is the probability of raising on L, given raising on the last two times that a low card was drawn, and the other three conditional probabilities are defined similarly. Sign tests failed to reject either of the two null hypotheses in each of the three conditions. We conclude from these analyses that in deciding whether to raise or drop, given a low card, the subject was most affected by his decision on the previous time a low card was drawn. Decisions on earlier periods had little or no effect on his current decision. These results differ from those obtained in two-person zerosum games in strategic form, where strong evidence has been found (e.g., Rapoport & Budescu, 1992, in press) for second- and even higher-order sequential dependencies. Table 4 presents the mean values of $p(R|R)$, $p(R|D)$, $p(R|R, R)$, $p(R|R, D)$, $p(R|D, R)$, and $p(R|D, D)$ by condition.

Similar analyses of first- and second-order conditional probabilities of calling were conducted for Player B. Denote by $p(C|C)$ the probability of calling on period $t + 1$, given a call on period t, and by $p(C|F)$ the probability of calling on period $t + 1$, given a fold decision on period t. Randomization of pure strategies by Player B implies no first-order dependencies, i.e., $p(C|C) = p(C|F)$. To test this null hypothesis, we computed for each type B player these two conditional probabilities and then subjected them to a sign test. The results show that $p(C|F) > p(C|C)$ for 16, 14, and 15 players in Conditions C2, C4, and C6, respectively. These results are statistically significant ($p < .05$) for Conditions C2 and C6, and only marginally significant ($p < .055$ for a one-sided test) for Condition C4. We conclude from these results that, like the type A players, most of the type B players produced non-Bernoulli type series exhibiting the overalternation bias.

We repeated the analysis for second-order effects, testing for each subject two null hypotheses, namely, $p(C|C, C) = p(C|C, F)$ and $p(C|F, C) = p(C|F, F)$. The sign test results were significant in only one of the six tests ($p(C|F, F) > p(C|F, C)$ in Condition C2); the remaining five tests failed to reject the null hypotheses. Similar to the players of type A, type B players seem to be mostly affected by their own decision on the immediately preceding period. We find no evidence for sequential biases based on a memory of more than a single period.

Table 5 presents the mean values of $p(C|C)$, $p(C|F)$, $p(C|C, C)$, $p(C|C, F)$, $p(C|F, C)$, and $p(C|F, F)$. Although the sign tests for the equality of first-order conditional probability yielded significant results, Table 5 shows that the mean difference between $p(C|C)$ and $p(C|F)$ is rather small.

A summary of the analyses reported above yields several behavioral regularities; we shall try to account for some of them in the subsequent section. As predicted by the equilibrium solution, players do not choose dominated strategies. As in two-person zerosum games presented in strategic form, both types of players mix their strategies, but the observed individual mixed strategies computed across all trials do not conform in general to the prediction (under the assumption of risk neutrality). In particular, the estimated individual probability of Player A raising on L does not increase with the size of the stake as predicted. Type B players tend to call more often than predicted, and type A players do not raise on low cards as often as they should. However, the results of most of the dyads in the moderate and

TABLE 4

Mean of First- and Second-Order Conditional Probabilities of Raising (Player A) by Condition

Condition	$p(R\|R)$	$p(R\|D)$	$p(R\|R, R)$	$p(R\|R, D)$	$p(R\|D, R)$	$p(R\|D, D)$
C2	0.35	0.57	0.34	0.34	0.50	0.63
C4	0.42	0.57	0.38	0.45	0.57	0.60
C6	0.38	0.51	0.46	0.43	0.55	0.53

TABLE 5

Means of First- and Second-Order Conditional Probabilities of Calling (Player B) by Condition

Condition	$p(C\|C)$	$p(C\|F)$	$p(C\|C, C)$	$p(C\|C, F)$	$p(C\|F, C)$	$p(C\|F, F)$
C2	0.66	0.78	0.64	0.70	0.75	0.88
C4	0.53	0.55	0.54	0.52	0.53	0.57
C6	0.47	0.53	0.50	0.49	0.49	0.57

high stake conditions can be accounted for by equilibria that correspond to different utility function pairs and give distinct advantage to the informed player (A). As in two-person zerosum games in strategic form, both types of players exhibit the sequential dependency known as the overalternation bias. The magnitude of this bias is stronger in the extensive than in the strategic form game. Because the strategic form of the game in Condition C2 gives rise to the "matching pennies" game, a direct comparison between the extensive and strategic forms is possible. A comparison of the results of Rapoport and Budescu (1992), who presented the "matching pennies" game in strategic form, and the results of Condition C2 shows a stronger overalternation bias in the latter study. In Condition C2, across all players of type A, $p(\mathrm{R}|\mathrm{R}) = .35 < p(\mathrm{R}|\mathrm{D}) = .57$, and across all players of type B, $p(\mathrm{C}|\mathrm{C}) = 0.66 < p(\mathrm{C}|\mathrm{F}) = .78$. In comparison, Rapoport and Budescu reported a considerably weaker overalternation bias with $p(i|i) = .465 < p(i|j) = .535$ across both the row and column players.

Adaptive Learning

When a game is played repeatedly, observed behavior is likely to change as the players acquire experience with the structure of the game and the behavior of other players. It is, therefore, natural to ask how well we can explain the dynamics of play—on both the aggregate and individual levels—in the iterated Poker game. To answer this question, we focus on the adaptive learning model of Roth and Erev (1995), which has been quite successful in tracking the dynamics of play in several different iterated two-person zerosum games in strategic form (Erev & Roth, 1995). In particular, Erev and Roth have shown that with the exception of sequential dependencies discussed above, the model can account for the major behavioral regularities observed in two-person zerosum games with a unique mixed-strategy equilibrium. Moreover, the model reproduced the major trends in the experimental results both when they were consistent with the equilibrium predictions and when they were not. The present study extends this model to our extensive form Poker game and examines the implications of sequential dependencies to reinforcement learning.

The Roth/Erev model is basically a reinforcement-based theory that models players in interactive decision making situations as cognitively-limited adaptive learners. The central assumption of this model is that any robustly descriptive model of human learning in strategic environments must be consistent with the characteristics of human and animal learning behavior described in the psychological learning literature. The model incorporates two basic principles of learning behavior—Thorndike's Law of Effect formulated about 100 years ago and Blackburn's Power Law of Practice formulated more than 60 years ago. It also incorporates two additional and extensively studied features of individual learning—generalization and recency (see Roth & Erev, 1995, for more details).

Specifically, the Roth/Erev learning model assumes that when the experiment commences, players start with initial "propensities" that, in turn, determine the probabilities with which they play their pure strategies. Presumably, these initial propensities reflect the players' beliefs based on past experience with the same or similar tasks (e.g., knowledge of Poker or, more generally, "bluffing" in two-person interactions). These probabilities are then updated over time in response to realized payoffs, with higher payoffs yielding larger changes in the probabilities. The basic model only applies to games without negative payoffs. The more general model (Erev & Roth, 1995), which we outline below, allows for both positive and negative payoffs. It assumes that at time $t = 1$, before any experience with the game has been acquired, each player n (n = A or B in our present study) has an initial propensity to play his kth pure strategy, given by some number $q_{\mathrm{nk}}(1)$. If player n plays his kth strategy at time t ($t = 1, 2, \ldots$) and receives a payoff of y, then the propensity to play any strategy j is updated by setting

$$q_{\mathrm{nj}}(t + 1) = \max[\nu, (1 - \phi)q_{\mathrm{nj}}(t) + \mathrm{E}_{\mathrm{nk(}}j, \mathrm{R}_t(y))]. \quad (4)$$

In Eq. (4), ϕ is a forgetting (or recency) parameter, which slowly reduces the importance of past experience; R is a reinforcement function translating payoffs into rewards; E is a generalization function determining how the experience of playing the kth strategy and receiving the reward $\mathrm{R}(y)$ is generalized to update each strategy j; and $\nu > 0$ is a technical parameter introduced to ensure that the propensities remain positive. The probability $p_{\mathrm{nk}}(t)$ that player n plays his kth strategy at time t is then computed from

$$p_{\mathrm{nk}}(t) = q_{\mathrm{nk}}(t)/\Sigma q_{\mathrm{nj}}(t), \quad (5)$$

where summation is taken over all of player n's pure strategies j.

The reinforcement and generalization functions in Eq. (4) are assumed to reflect structural properties of the game as well as empirical observations. Following the work of Herrnstein (1961), who demonstrated a linear relation between reinforcement and choice probabilities, the reinforcement function in the learning model assumes the form $\mathrm{R}_t(y) = y - \rho(t)$, where $\rho(t)$ is a reference point parameter depending on period t. Erev and Roth denote the reference point at period 1 by $\rho(1)$, and then assume a linear weighing function for updating it:

$$\rho(t+1) = \begin{cases} (1 - w^{+})\rho(t), & \text{if } y \geq \rho(t) \\ (1 - w^{-})\rho(t), & \text{if } y < \rho(t), \end{cases}$$

where w^{+} and w^{-} are the weights assigned to positive and negative reinforcement (relative to $\rho(t)$), respectively.

In the present study, where the (two) strategies cannot be ordered along a natural dimension determining "strategy similarity", Roth and Erev's generalization function takes the form:

$$E_{nk}(j, R(y)) = \begin{cases} R(y)(1 - \epsilon), & \text{if } j = k, \\ R(y)\epsilon/(m - 1), & \text{otherwise}, \end{cases}$$

where m is the number of pure strategies.

We consider below two alternative approaches to estimate the initial propensities. Under the first approach, the initial propensities are estimated from the initial play of the game. Specifically, note that Eq. (5) implies that

$$q_{nk}(1) = p_{nk}(1)S(1),$$

where $S(1) = \Sigma q_{nk}(1)$ is a free parameter reflecting the "strength" of the initial propensities. In applying the model, we fix S(1) and determine $p_{nk}(1)$ by pooling the choices of all the players with the same strategy sets and then assigning to each of them initial probabilities equal to the observed mean probabilities in the first block of 24 periods. Our second approach incorporates a weaker assumption of uniformly distributed initial propensities.

The present analysis is not intended to test the model in its generality but, rather, to examine if the learning model with the same parameter values that have been found to provide the best approximation of behavior in strategic form games (Erev & Roth, 1995) provides a reasonable description of the behavioral regularities observed in the present study. Therefore, rather than attempting to estimate parameter values through a search in a seven-dimensional space, we have set all the parameters at the values reported by Erev and Roth to provide the best fit in their study of adaptive learning in several iterated games in strategic form, namely, $\epsilon = 0.2$, $\nu = 0.0001$, $\phi = 0.001$, $S(1) = 3$ (mean absolute payoff), $\rho(0) = 0$, $w^{+} = 0.01$, and $w^{-} = 0.02$.

Before applying the model to our data, it is instructive to consider the implications of the sequential dependencies found in our study for the model's potential usefulness. Under the cognitive approach of Roth\Erev, these dependencies can result from "sequential cognitive strategies." Two types of robust sequential strategies have been identified in previous research:

(a) "Alternation (ALT) and repetition strategies"—strategies that condition the *decision* on period t on the previous decision. Simple alternation strategies appear to be used even by rats. As revealed, for example, by Rapoport and Budescu (1992), longer patterns are observed in certain settings.

(b) "Gambler fallacy (GF) strategies"—strategies that condition the decision on period t on the *outcome* of the previous period. A strategy of this type involves the decision to gamble following a loss (in the expectation that luck will "correct itself"). This type of strategy leads to behavior that contradicts the behavioral interpretation of the Law of Effect (see, e.g., Lee, 1971).

To evaluate the effect of sequential strategies, we derived the model's predictions (by simulation) either with or without such strategies. The two basic sequential strategies—(a) and (b) above—were considered in each of the three experimental conditions.

Figure 2 portrays the simulated learning functions for each of the three experimental conditions C2, C4, and C6. In each condition, the top row portrays the simulation results for the case where learning is assumed to take place on each period, whereas the bottom row portrays the simulation results for the case where the propensities change only on periods in which a decision was actually made by one or both of the players. Initial propensities for all the simulations in Fig. 2 have been estimated under the assumption (our second estimation procedure) of a uniform distribution.

In each individual plot in Fig. 2 the probability of call, q, is displayed on the horizontal axis, and the probability of raise on L (bluff), p, on the vertical axis. The simulation begins at the joint initial probabilities, which are designated by an empty triangle in each plot, and ends (after 10,000 periods) in the filled circle. The equilibrium solution in mixed strategies (for risk neutral players) is depicted by an empty circle. The differences between the five simulations in each row of the figure are entirely due different random processes. Specifically, we use the same uniformly distributed initial propensities in each row and column. The differences in the starting points of the simulations are due to the different probabilities of card drawing and choice in the first 25 trials.

Figure 3 displays similar plots of simulated learning functions, generated in exactly the same way, for the

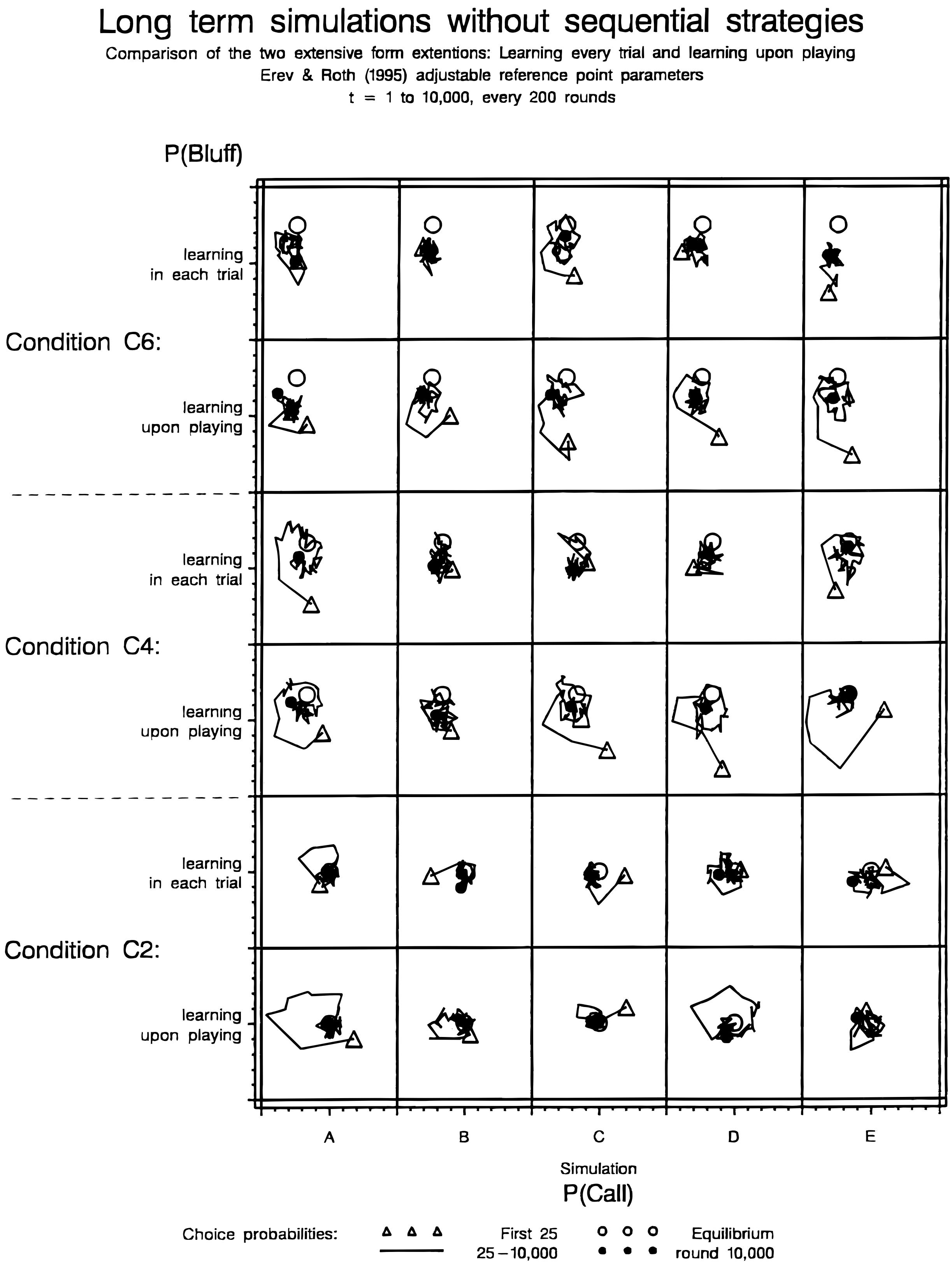

FIG. 2. Long-term simulations of the Roth/Erev model with no sequential dependencies.

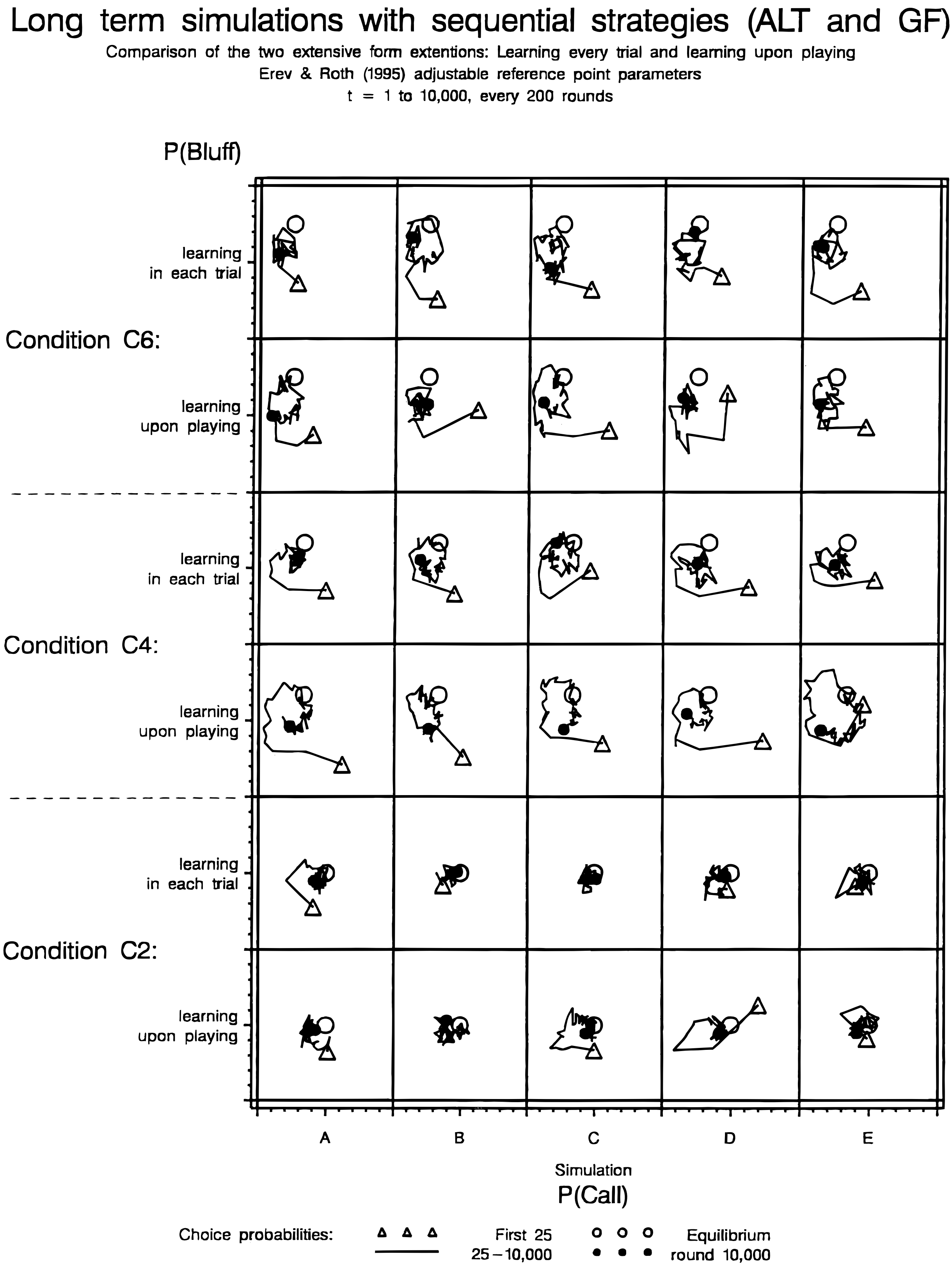

FIG. 3. Long-term simulations of the Roth/Erev model with sequential dependencies.

version of the learning model incorporating both sequential dependencies. In these simulations, each virtual player had four strategies: the two nondominated stage game strategies and the two repeated game strategies described above (ALT and GF). Similarly to the simulation results displayed in Fig. 2, players started with uniform initials among the different strategies. Note that the actual choice, given a selection of a specific strategy, is a function of the strategy and the relevant recent event. For example, Player A who chose the alternation (ALT) strategy on trial *t* dropped (on Low card) only if she raised on the previous relevant trial.

Comparison of Figs. 2 and 3 reveals that the model's simulations are relatively unaffected by the assumed utilization of the two sequential strategies. The simulations imply three major predictions independently of the strategy space. First, with only a few exceptions, learning proceeds in a clockwise fashion regardless of the values of the initial propensities. Second, learning is relatively fast in the first 200 periods (depicted in each plot by a straight line emanating from the empty triangle), and then slows down considerably. Third, although best reply is not assumed by the Roth/Erev model, the simulated joint probabilities (p, q) are seen to converge (rather slowly) to the equilibrium solution in all cases in Conditions C2 and C4 but only approach it in Condition C6.

Comparison of the two sets of five simulations within each condition shows that the assumption about the timing of learning matters. When Player B learns only on periods in which she is called upon to make a decision, the initial periods are played at the bottom right side of the equilibrium, i.e., Player B calls "too often" and player A does not raise as frequently as he should. This tendency, that corresponds to the experimental results, is weakened when player B is assumed to learn on periods on which she is not called upon to play. This result suggests that under the present learning model the "learning upon playing" assumption can account for the deviation of the observed results from equilibrium.

The robustness of the model's predictions with respect to the utilization of sequential dependencies has both encouraging and discouraging implications. On the negative side, this result implies that the existence of sequential dependencies cannot be used to reject the reinforcement-based learning model in its full generality. Rather, the existence of sequential dependencies can be used to reject specific models (e.g., the two-strategy model predicting no sequential dependencies). Because a variety of sequential strategies can be used by the subjects, and because these strategies are not easily observable, this conclusion implies that the present approach cannot lead to a precise prediction of behavior.

On the more favorable side, the present results imply that the model can be used to derive (by simulation) testable, nontrivial, and robust predictions even without precise understanding of the nature of the sequential strategies that subjects use. In fact, in the present context, the sequential dependencies can be ignored without affecting the derived predictions. The analysis presented below proceeds under this assumption with the two-strategy model being used to derive the predictions.

Figure 4 displays the observed and simulated probabilities p and q for the aggregate data. The observed results are averaged over all the twenty players in each condition. Each row in the figure corresponds to a different experimental condition. The three (Condition C2) or five (Conditions C4 and C6) points in each plot correspond to the mean joint probabilities of bluff and call in blocks of 24 periods. Column 1 of Figure 4 shows the observed mean probabilities of call (square) and bluff (triangle). Column 2 shows the simulated results for the case of learning upon playing, whereas column 3 shows the simulation results, also in blocks of 24 periods, for learning on each period. The isolated points are the equilibrium solutions for risk-neutral players (Eq. 3). All the simulations were initialized from the first block of the experimental results; the initial propensities were assessed from the actual choice probabilities.

Inspection of Fig. 4 shows that the adaptive learning model for the case of learning upon playing tracks the observed mean results quite accurately. There is a small down trend over blocks in the data, which is captured by the first set of simulated results (column 2). The second set of simulations (column 3) does considerably worse.

As a second test of the learning model, we computed the first-order conditional probabilities from the simulated data. For this computation, we conducted 100 simulations, each for 100 periods. The mean first-order conditional probabilities are presented in Table 6. The results are shown separately for type A and B players, with or without sequential dependencies, assuming either learning on each period or learning upon playing.

Table 6 shows that the inequality $p(\text{R|D}) > p(\text{R|R})$, which was observed for almost all the type A players, (see columns 2 and 3 of Table 4) holds in the version of the model which includes sequential strategies. Similarly, the inequality $p(\text{C|F}) > p(\text{C|C})$, which was observed for most of the type B players (see columns 2 and 3 of Table 5), only holds when sequential strategies are assumed by the model. The conditional probabilities for type A players are slightly higher when learning is assumed to take place on every period, whereas the opposite results obtain for the type B players. In correspondence with the observed results, the difference

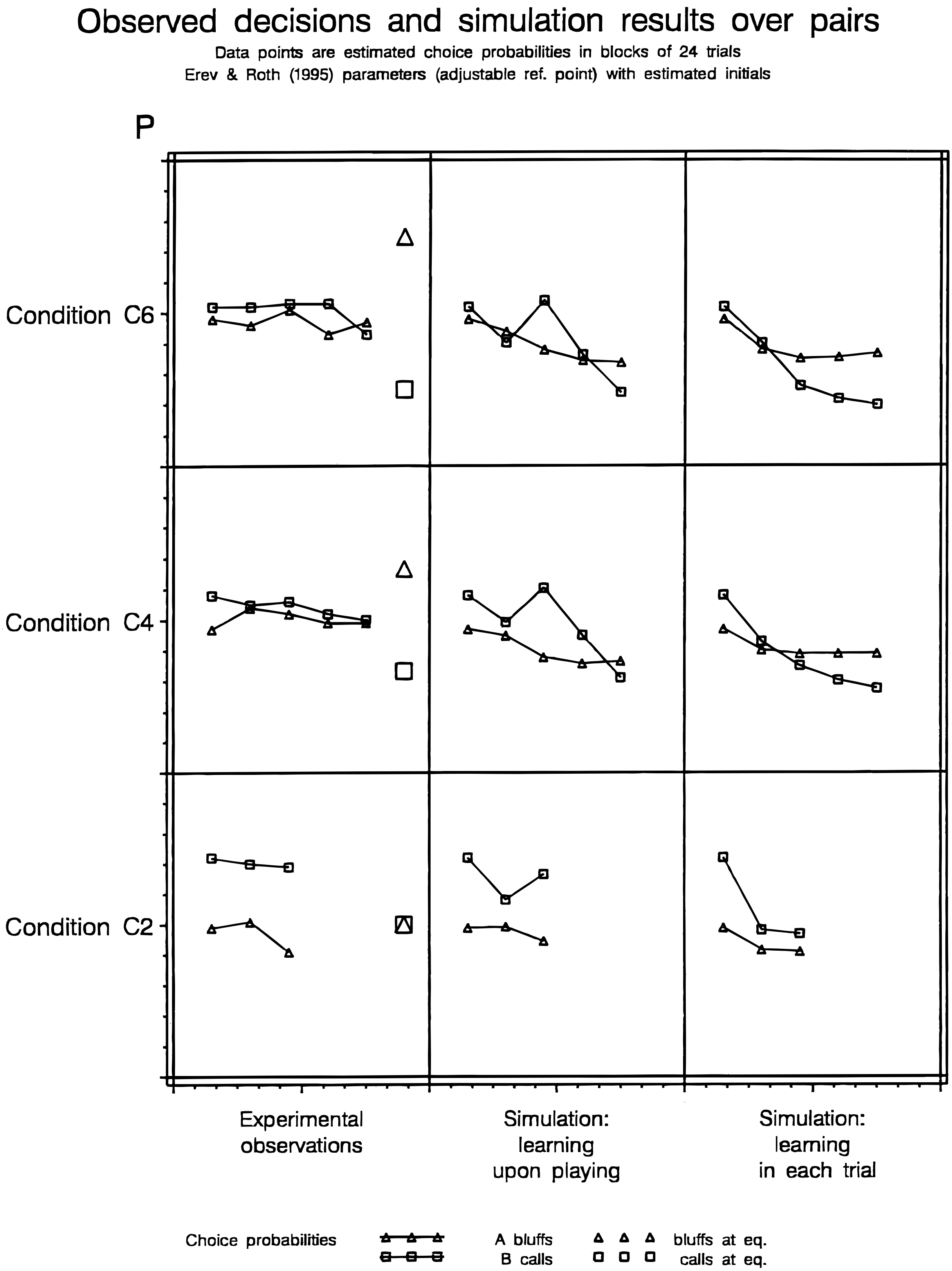

FIG. 4. Mean joint probabilities of bluff and call: observed and simulation results.

TABLE 6

Mean First-Order Conditional Probabilities of Raising (Player A) and Calling (Player B) in the First 100 Rounds of Simulation

Sequential strategy	Timing of learning	Condition	p(R|R)	p(R|D)	p(C|C)	p(C|F)
Yes	Upon Playing	C2	.26	.74	.49	.66
		C4	.17	.66	.40	.42
		C6	.19	.72	.35	.41
	On Every Period	C2	.25	.75	.35	.53
		C4	.24	.74	.30	.38
		C6	.22	.70	.23	.29
No	Upon Playing	C2	.40	.35	.37	.36
		C4	.68	.48	.59	.31
		C6	.34	.32	.54	.27
	On Every Period	C2	.48	.57	.64	.38
		C4	.43	.41	.50	.23
		C6	.43	.46	.43	.17

p(R|D) − p(R|R) for type A players far exceeds the difference p(C|F) − p(C|C) for type B players in all three conditions.

We turn next from aggregate to individual results. Figure 5 shows the observed mean joint probabilities p and q for each dyad separately. As in the previous figure, each mean is based on the choices in blocks of 24 periods. The results for the individual dyads in Conditions C2, C4, and C6 are portrayed in Figs. 5A, 5B, and 5C, respectively.

At a first glance, the plots in Fig. 5 do not appear to conform to the model's clockwise circle-in prediction. Many of the dyads "move" in different directions. However, clockwise behavior is predicted for a long sequence whereas the observed results are based on short sequences of length 72 or 120 periods. Moreover, because of the probabilistic nature of the model, violations may be expected.[4]

To compare the observed and simulated deviations from the model's clockwise prediction, we conducted the following two-stage analysis. First, a simulation was run (using the two-strategy learning upon playing model) to derive the model's prediction for each of the 60 experimental dyads. Each simulation was initialized from the relevant dyad choice probability in the first half of the experiment. The simulations were then conducted on the second half of the experiment (namely, periods 37–72 in Condition C2 and periods 71–120 in Conditions C4 and C6).

In the second part of the analysis, we computed the proportion of simulated and experimental subjects that moved in the predicted clockwise direction. About 62% of type A players and 61% of type B players moved in the predicted direction. These values are significantly higher than the chance prediction of 50% (sign test, $p < .05$). The corresponding results for the simulated players were 64% for type A and 60% for type B players; these percentages do not differ significantly from the ones computed for the experimental subjects.

[4] It should be noted that the fact that the plots in Fig. 5 appear to be less orderly than the ones depicted in Figs. 2 and 3 does not imply a real difference between the simulated and experimental results, only a difference in group size. Each point in Figs. 2 and 3 is based on 200 observations, whereas each point in Fig. 5 is based on 24 observations only.

CONCLUSIONS

Some of the regularities that characterize interactive behavior in two-person zerosum games in strategic form with a unique equilibrium in mixed strategies no longer hold when the game is presented in extensive form with asymmetry in information. In particular, the aggregate results no longer approximate the equilibrium solution, the first-order sequential dependency known as the overalternation bias is significantly enhanced, whereas higher-order sequential dependencies weaken or completely disappear. Compared to the mixed-strategy equilibrium for risk-neutral players, the informed player in the Poker game does not raise as often as predicted, whereas the uninformed player calls too often. However, the aggregate results are not necessarily representative of the individual results. There are large individual differences in the propensity to raise and in the propensity to call, which far exceed the individual differences observed in two-person zerosum games in strategic form with either symmetric or asymmetric players.

Based on the parameter values estimated by Erev and Roth and the auxiliary assumption that learning takes place only on periods when players are called to act, the learning model captures several of the behavioral regularities observed in the study. In particular, the simulated players deviated from the equilibrium prediction in the same direction as the experimental subjects, namely, "too many calls and not enough bluffs." These deviations were observed even when the simulations were initialized by uniform initial propensities (Figs. 2 and 3). In addition, the adaptive learning model successfully predicted the average dynamic (clockwise trend) as well as the percentage of violations of this prediction. About 62 percent of the experimental and simulated subjects exhibited the predicted dynamic. When sequential strategies are assumed, the

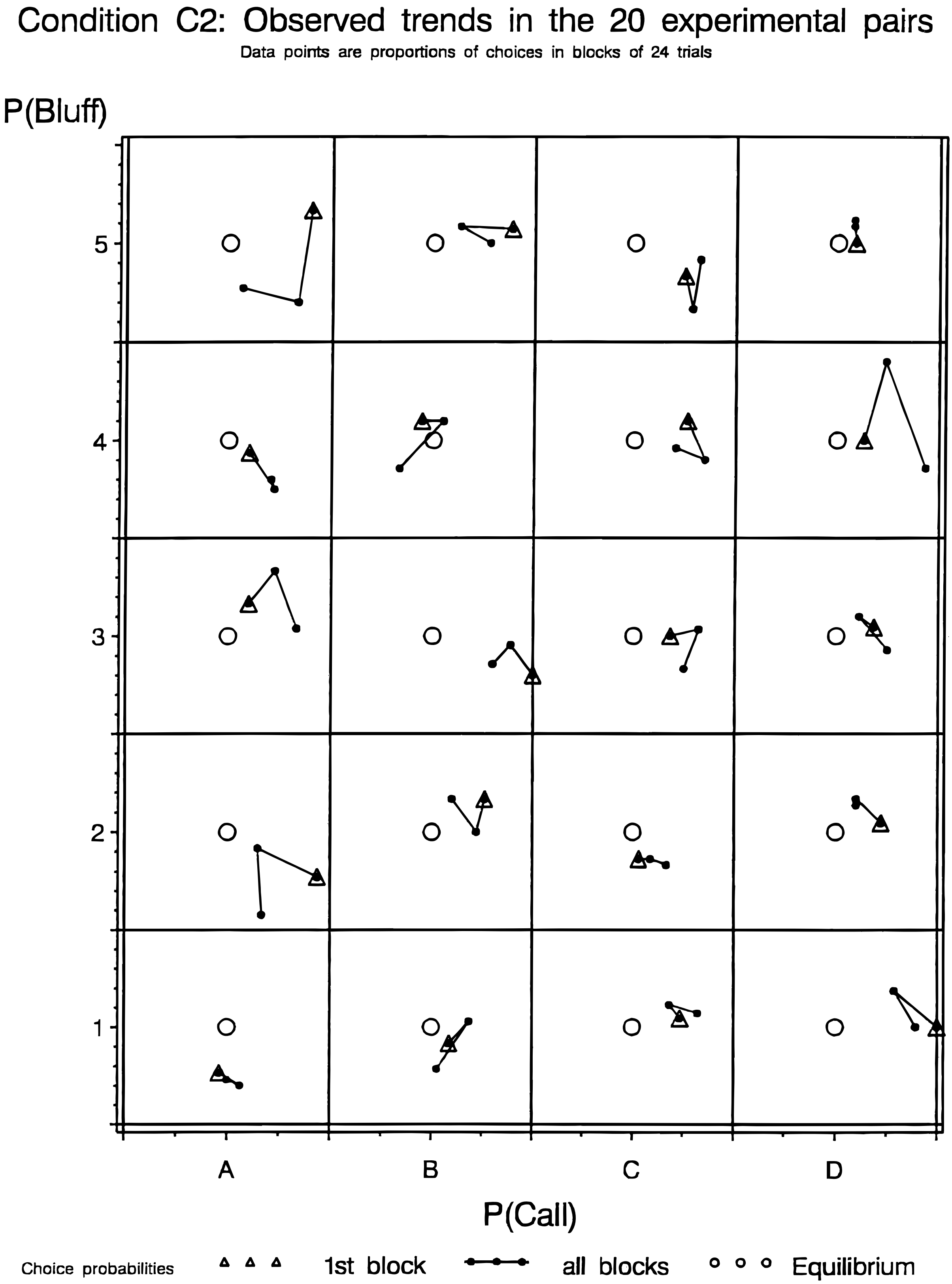

FIG. 5. Individual joint probabilities of bluff and call by condition and block.

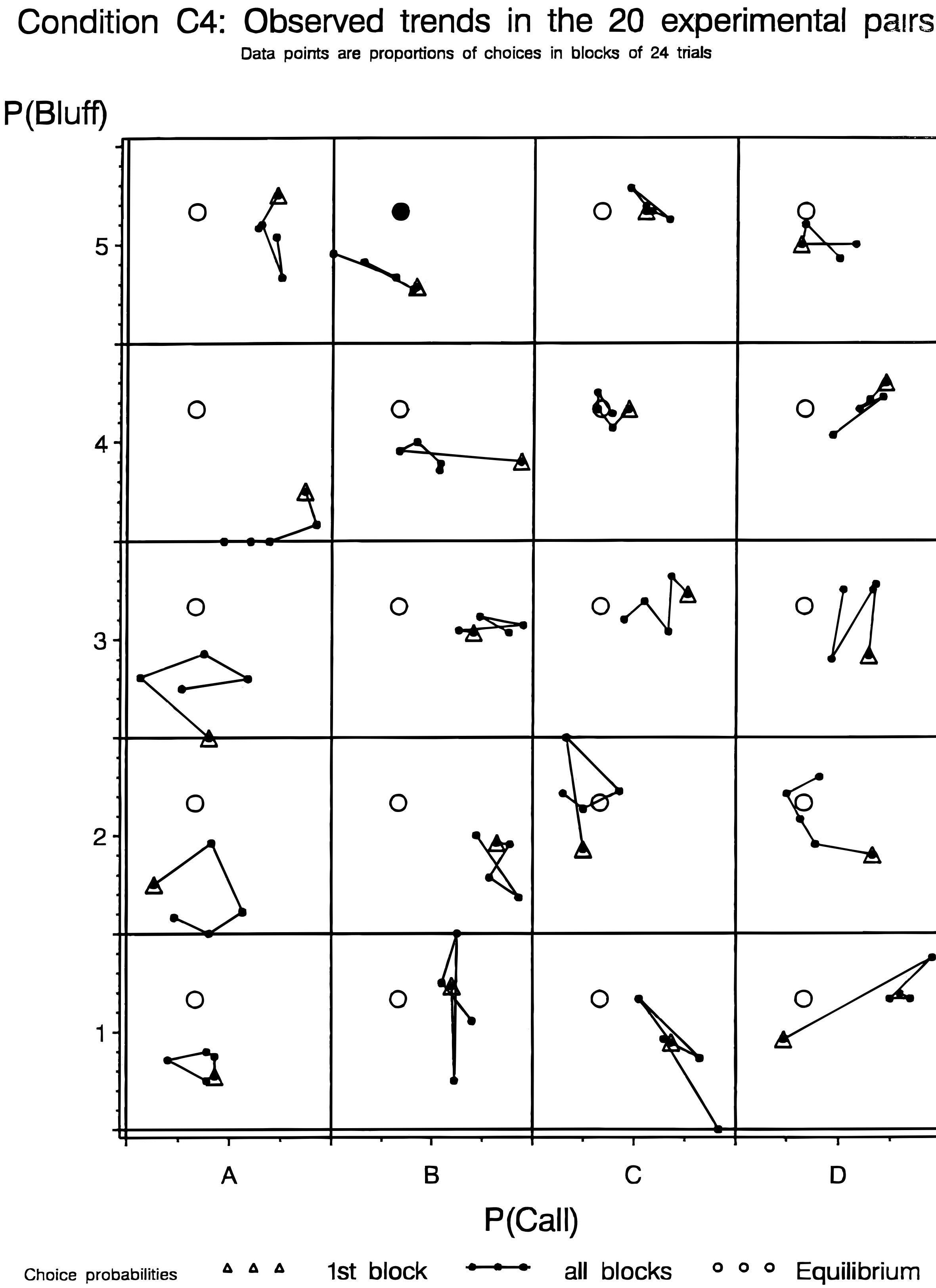

FIG. 5—*Continued*

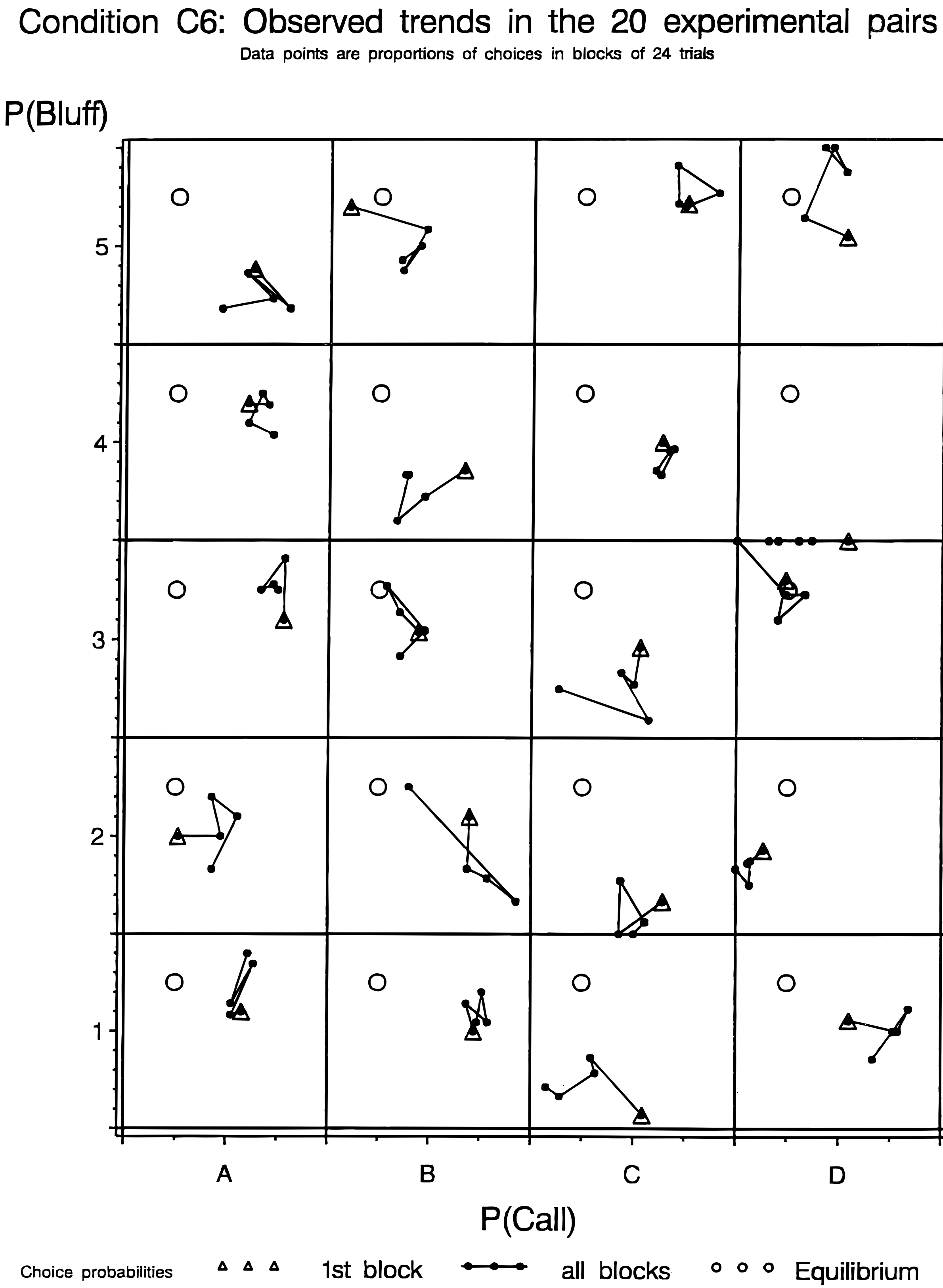

FIG. 5—*Continued*

model accounted for the sequential dependencies observed in the data. In line with the experimental results, the overalternation bias exhibited by type A players was stronger and less sensitive to the experimental condition than the overalternation bias exhibited by the type B players.

The fact that the Roth/Erev model provided a reasonably good first-order approximation to the major regularities found in both extensive and strategic games supports the hypothesis that the same psychological processes underlie decisions in both cases. Only one modeling decision had to be made in extending the model from strategic to extensive form games (learning upon playing), and the present results give a clear answer as to the more appropriate modeling. The results reported above suggest that the model's descriptive power is maximized when players are assumed to learn only when required to act. Further studies using different extensive form games are required to test this hypothesis.

REFERENCES

Binmore, K. (1992). *Fun and games: A text on game theory.* Lexington, MA: D. C. Heath.

Binmore, K., Swierzbinski, J., & Proulx, C. (1996). *Does Minimax work? An experimental study.* Paper presented at the Annual Conference of the Public Choice Society, Houston, Tx.

Brown, J. N. & Rosenthal, R. (1990). Testing the minimax hypothesis: A reexamination of O'Neill's game experiment. *Econometrica,* **58,** 1065–1081.

Budescu, D. V., & Rapoport, A. (1994). Subjective randomization in one- and two-person games. *Journal of Behavioral Decision Making,* **7,** 261–278.

Erev, I., & Roth, A. E. (1995). *On the need for low rationality, cognitive game theory: Reinforcement learning in experimental games with unique, mixed strategy equilibria.* Unpublished manuscript, Department of Economics, University of Pittsburgh.

Fox, J. (1972). The learning strategies in simple two-person games without saddle-point. *Behavioral Science,* **17,** 300–308.

Fudenberg, D. & Kreps, D. (1993). Learning mixed equilibria. *Games and Economic Behavior,* **5,** 320–367.

Greenberg, J. (1994). *Social situations without commonality of beliefs: Worlds apart—but acting together.* Unpublished manuscript, Department of Economics, McGill University.

Güth, W. Huck, S. & Rapoport, A. (in press). The limits on the positional order effect: Can it support silent threats and non-equilibrium behavior? *Journal of Economic Behavior and Organization.*

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory,* **2,** 1–23.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior,* **13,** 243–266.

Received: April 22, 1996

Kreps, D. (1990). *Game theory and economic modeling.* Oxford: Oxford Univ. Press.

Lee, W. (1971). *Decision theory and human behavior.* New York: Wiley.

Lieberman, B. (1960). Human behavior in a strictly determined 3 × 3 matrix game. *Behavioral Science,* **5,** 317–322.

Lieberman, B. (1962). Experimental studies of conflict in some two- and three-person games. In J. H. Criswell, H. Solomon, & P. Suppes (Eds.), *Mathematical models in small group processes* (pp. 203–220), Stanford, CA: Stanford Univ. Press.

Messick, D. M. (1967). Interdependent decision strategies in zero-sum games: A computer-controlled study. *Behavioral Science,* **12,** 33–48.

Mookherjee, D. & Sopher, B. (1994). Learning behavior in an experimental matching pennies game. *Games and Economic Behavior,* **7,** 62–91.

Morin, R. E. (1960). Strategies in games with saddlepoints. *Psychological Reports,* **7,** 479–485.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences,* **84,** 2106–2109.

Osborne, M. J. & Rubinstein, A. (1994). *A course in game theory.* Cambridge, MA: MIT Press.

Radner, R. & Rosenthal, R. W. (1982). Private information and pure-strategy equilibria. *Mathematics of Operations Research,* **7,** 401–409.

Rapoport, A. (in press). Order of play in strategically equivalent games in extensive form. *International Journal of Game Theory.*

Rapoport, A. & Boebel, R. B. (1992). Mixed strategies in strictly competitive games: A further test of the minimax hypothesis. *Games and Economic Behavior,* **4,** 261–283.

Rapoport, A. & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General,* **121,** 352–363.

Rapoport, A. & Budescu, D. V. (in press). Randomization in individual choice behavior. *Psychological Review.*

Rapoport, A. & Fuller, M. A. (1995). *Coordination in noncooperative three-person games under different information structures.* Unpublished manuscript.

Roth, A. E. & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior,* **8,** 164–212.

Rubinstein, A. (1991). Comments on the interpretation of game theory. *Econometrica,* **59,** 257–274.

Shachat, J. M. (1995). *Mixed strategy play and the Minimax hypothesis.* Unpublished manuscript, Department of Economics, University of Arizona.

Stahl, D. (1988). On the instability of mixed strategy Nash equilibria. *Journal of Economic Behavior and Organization,* **9,** 59–69.

Suppes, P. & Atkinson, R. C. (1960). *Markov learning models for multiperson interactions.* Stanford, CA: Stanford Univ. Press.

Tversky, A. & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty,* **5,** 297–323.

Von Neumann, J. & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton: Princeton Univ. Press.