# The effect of matching mechanism on learning in games played under limited information


Nick Feltovich[1]
Department of Economics
Monash University
Clayton, VIC 3800, Australia
FAX +61 3 9905 5476
nicholas.feltovich@monash.edu

Sobei H. Oda
Faculty of Economics
Kyoto Sangyo University
Kamigamo, Kita-ku, Kyoto 603-8555, Japan
FAX +81 75 7051949
oda@cc.kyoto-su.ac.jp



Abstract: We examine the effect of the matching mechanism on learning in 2x2 games. Six games are played repeatedly under either *fixed pairs* or *random matching*. Unlike most economics experiments, the games are played under *limited information*: subjects are never shown the games' payoff matrices nor given information about opponent payoffs. We find that behavior, while initially similar between treatments, diverges over time. In most but not all games, fixed-pairs matching is associated with increased coordination on pure-strategy Nash equilibria, higher-payoff equilibria being reached, and faster convergence toward pure-strategy play.

## 1. Introduction

Colin Camerer (2003), summarizing the state of experimental game theory research in his behavioral game theory text, said, "[t]here are no interesting games in which subjects reach a predicted equilibrium immediately. And there are no games so complicated that subjects do not converge in the direction of equilibrium (perhaps quite close to it) with enough experience" (p. 20). The implication of this passage is that, while game-theoretic concepts like Nash equilibrium are useful for characterizing how individuals behave after acquiring sufficient experience, a true behavioral game theory must incorporate a description of how individuals learn.[2] Though the beginnings of a learning theory are well established (Camerer and Ho, 1999; Fudenberg and Levine, 1998; Roth and Erev, 1995), we are far from a consensus regarding the exact nature of learning in games. In order for progress toward some kind of consensus to continue, more detailed study must be made of factors that can influence how decision makers learn.

The objective of this paper is to take a small step in this direction. We concentrate on the effects of one particular experimental design manipulation—the protocol used to match players to opponents—on the manner in which behavior changes over time. Many economics experiments comprise more rounds than there are potential opponents. For such experiments, there are two commonly-used matching mechanisms. Under *fixed pairs*, players are matched repeatedly to the same opponent. Under *random matching*, opponents are randomly assigned in every round. We conduct a laboratory experiment in which subjects repeatedly play several two-player games under one of these mechanisms. In order to focus on learning, rather than other phenomena that might be sensitive to the matching protocol, as well as to minimize effects arising from other-regarding preferences, we give subjects only limited information about the games—in contrast to most previous studies of matching mechanisms, which have focused on behavior under complete information (see Section 2.1). In our design, subjects are told they are playing a game, but are not given any information about payoffs before playing, and while they can learn about their own payoffs via end-of-round feedback, they never receive information about opponent payoffs; thus, even subjects who figure out the payoff matrix are unlikely to have much confidence that their opponents have also figured it out (let alone the common-knowledge assumption).

Our results suggest that the matching mechanism can indeed have a sizable effect on behavior in general, and learning in particular. Levels of cooperation are frequently higher under fixed pairs than under random matching. Also, convergence to equilibrium is usually faster under fixed pairs, though in one game, the opposite is true. The matching mechanism also can have an effect on the likelihood of behavior converging to one equilibrium versus another: in one game, fixed-pairs matching actually increases the chance of players getting "stuck" on a low-payoff equilibrium, while in the others, fixed pairs makes the high-payoff equilibrium more likely.

## 2. The experiment

Figure 1 shows the games used in the experiment.

[Figure 1 about here]

While each game is symmetric and 2x2, they differ in some important ways. Prisoners' Dilemma has a strictly dominant strategy, D, and thus a unique Nash equilibrium. The other games have multiple Nash equilibria and no dominant strategies. In Battle of the Sexes, the strategies are strategic substitutes (a player's strategy becomes less attractive, the more likely the opposing player is to choose it), so that the two pure-strategy equilibria are asymmetric; in the other four games with multiple equilibria, the strategies are strategic complements (a strategy becomes more attractive as the likelihood of the opponent playing it increases), so their pure-strategy equilibria are symmetric. For ease of exposition, we have ordered the two strategies in each game in such a way that the first strategy is "nice" in the sense that it tends to be associated with higher payoffs for the other player than the second strategy; we have also labeled the actions C (for "cooperate") and D (for "defect"), though these terms are more literally meaningful in some games than others.

Because each game has only two actions, a player's strategy can be characterized by the associated probability of choosing C. Thus, any *strategy pair* (one strategy for each player) can be written in the form (Prob(Row player chooses C), Prob(Column player chooses C)). Table 1 shows the games' equilibrium strategy pairs, along with the implied frequencies of C choices; these probabilities will serve as useful benchmarks for our results, even though we are not explicitly testing equilibrium predictions in this paper. (Note that uniform random play—the (0.5, 0.5) strategy pair—is not an equilibrium of any of these games.)

[Table 1 about here]

Versions of these games have been used in many experiments. We do not attempt a review of all relevant literature here; much can be found in the sections on coordination games and social dilemmas in Camerer (2003) and Kagel and Roth (1995).

### 2.1 Experimental design and related literature

As already mentioned, our primary design variable is the matching treatment. In the *fixed-pairs* (F) treatment, subjects played all 40 rounds of a game against the same opponent, though opponents did change from game to game. In the *random-matching* (R) treatment, subjects were randomly assigned to opponents in each round, with every potential opponent (including the previous-round opponent) equally likely. Since the number of rounds of each game was more than the number of potential opponents in an experimental session, subjects in the R treatment faced the same opponent more than once in the same game; however, as no identifying information about opponents was given to subjects, they would never have been able to tell whether—and in which round—they had previously been matched to their current opponent.

There has been some work examining the effects of matching mechanisms, mostly using games played under complete information. Many such studies used social dilemmas, since under complete information, one might expect to see more cooperative behavior under fixed pairs—where incentives for reputation building are stronger—than under random matching, and that this effect would be stronger in early rounds than later ones (as the value of a reputation should decline as the number of remaining rounds becomes small). While Andreoni and Croson (2008) found no systematic difference between fixed pairs and random matching in their survey of early public-good experiments, other studies have typically found higher levels of cooperation under fixed pairs in the Prisoners' Dilemma (Ahn et al., 2001; Duffy and Ochs, 2008), and in other environments with features of social dilemmas (Charness and Garoupa, 2001; Huck et al., 2001). Similarly, studies using coordination games with Pareto-ranked pure-strategy equilibria can be Pareto-ranked have tended to find either no difference (Schmidt et al., 2003) or higher efficiency under fixed pairs than random matching (Van Huyck et al., 1990; Clark and Sefton, 2001). McKelvey and Palfrey (2001) examined eight simple games in their "Full Information" treatment, including coordination games and a version of the Prisoners' Dilemma, and found higher payoffs under fixed pairs than under random matching, though they didn't report aggregate choice frequencies, so it is not possible to determine how learning was affected. Danz et al. (2012), on the other hand, found no systematic difference in learning between fixed-pairs and random-matching treatments in a 3x3 game.

We stress that the results reported above came from experimental treatments with complete information about payoffs, making it difficult to disentangle differences in learning from differences in attempts at reputation building, or in the prevalence of early-round signaling. In contrast, our experiment uses a *limited-information* design, with no display of the payoff matrix to subjects—either publicly or privately. Instead, subjects receive some payoff information as part of their end-of-round feedback; specifically, they are informed after each round of their opponent's choice and their own payoff in the just-completed round. While this is enough information to allow subjects to piece together the relationship between outcomes and their own payoffs within a few rounds, it differs from the usual complete-information treatment in two notable ways. First, subjects never receive information about their opponents' payoffs; we believe it is reasonable to expect that the lack of this particular information should serve to undermine any effects on behavior of other-regarding preferences, even in games such as PD where one might normally expect such preferences to be present. Second, it is exceedingly unlikely that the structure of the game was *common knowledge* amongst the subjects; indeed, even a subject who has figured out a game's payoff structure would probably have little confidence that her opponent also will have done so, thus all but ruling out reputation building or early-round signaling.

There have been few experiments looking at different matching mechanisms under limited information, and their collective evidence of an effect is mixed. Cox et al. (2001) found that subjects with limited information about opponent payoffs were more likely to reach equilibrium in the actual game under fixed pairs than under random matching. McKelvey and Palfrey (2001), mentioned above, had a "No Information" treatment that was similar to ours except that subjects didn't receive information about opponent action choices; their reported results suggest that the matching mechanism didn't matter in more than half of their games.[3] Chen's (2003) comparison of pricing mechanisms for allocating shared resources within an organization also found no significant differences between fixed pairs and random matching.

While our interest is in the effects of the matching mechanism, limited-information settings have also been used to investigate other questions. A natural topic is how individuals learn how to behave in a setting where relevant information such as own or opponent payoffs is, at least initially, unknown, so that individuals are learning simultaneously about the rules of the game and about how they should play the game. There has been theoretical work on this topic (e.g., Dekel et al., 2004; Kaneko and Kline, 2008; Hanaki et al., 2009), as well as experiments. Overall, subjects seem to have difficulty learning initially-unknown aspects of games (Feltovich, 2000; Oechssler and Schipper, 2003), but nonetheless manage to figure out the equilibrium (Cox et al., 2001; Shachat and Walker, 2004).[4] The extent to which subjects can learn a game's payoff structure also depends on characteristics of the game such as uniqueness of equilibria and symmetry (Gerber, 2006), and on the amount of feedback given (Nicklisch, 2011).

### 2.2 Experimental procedures

Subjects in our experiment played all six games under either fixed pairs or random matching. We used a two-population design, with half of the subjects in each session designated "row players" and half "column players", and with subjects of one type only matched to the other type. The games were always ordered CG-SH-BoS-SH-PD-SH, but we varied the order of the three Stag Hunt games (SHH-SHM-SHL, SHM-SHL-SHH, and SHL-SHH-SHM) as a partial attempt to control for order effects. We also varied the order in which the actions appeared on subjects' screens: C-D (cooperative action on top or left) and D-C (cooperative action on bottom or right).

Our manipulations of game ordering, action ordering, and matching mechanism were between-subjects, while our manipulation of the game was within-subject.

Experimental sessions took place in 2003 and 2004 at the Kyoto Experimental Economics Laboratory (KEEL) at Kyoto Sangyo University. Subjects were primarily undergraduates, recruited via a database of participants in other experiments and via advertisements posted on campus. No one participated in more than one session. At the beginning of a session, each subject was seated and given a set of written instructions (an English translation is in Appendix A). After a few minutes, the instructions were read aloud by the monitor in an effort to make the rules common knowledge. Partitions prevented subjects from seeing others' computer screens, and subjects were asked not to communicate with each other. The experiment was programmed in the Japanese version of z-Tree (Fischbacher, 2007), and all interaction took place via the computer network. Subjects were asked not to write down any results or other information.

At the beginning of a session, subjects were told their type (row or column player). Prior to the first round of each game, they were reminded that they were beginning a new game. In every round, they were prompted to choose an action; these were given generic names (R1 and R2 for row players, C1 and C2 for column players). After all subjects had chosen, each was told her own action choice, her opponent's action choice, and her own payoff. In the R treatment, subjects were rematched after each round; in the F treatment, they were rematched only when the 40 rounds ended and a new game began. In both treatments, subjects were informed of the matching mechanism in both oral and written instructions.

At the end of an session, one round of one game was randomly chosen, and each subject was paid 200 Japanese yen (at the time of the experiment, equivalent to roughly USD 1.90) for each point earned in that round. In addition, subjects were paid a showup fee of 3000 yen, from which negative payoffs were subtracted, if necessary.

## 3. Experimental results

A total of 13 sessions were conducted: 6 of the R treatment and 7 of the F treatment. The number of subjects varied from 6 to 28 in the F sessions, and from 10 to 26 in the R sessions. Some session information is shown in Table B1 in Appendix B; the raw data are available from the corresponding author upon request.

### 3.1 Aggregate behavior

Table 2 reports C-choice frequencies in each game for rounds 1-5 as a proxy for initial behavior, rounds 36-40 as a proxy for endgame behavior, and over all rounds. In addition to these levels, the table shows the results of robust rank-order tests of significance of differences between the F and R treatments for each game and time period, using session-level data (see Siegel and Castellan, 1988, for descriptions of the nonparametric tests used in this paper, and see Feltovich, 2005, for critical values for the robust rank-order test).

[Table 2 about here]

In both treatments, frequencies of C choices begin at levels comparable to those implied by uniform random play, as one would expect since subjects initially have no information about payoffs. Indeed, for rounds 1-5, nonparametric Wilcoxon signed-ranks tests find significant differences from uniform random play (two-tailed test, p-values of 0.05 or lower) in only two of the six games in the F treatment and none in the R treatment, and we find no significant differences between the treatments. In later rounds, on the other hand, we find many differences, and when significant, they always point in the same direction: more C choices in the F treatment than in the R treatment. These differences are usually consistent with better outcomes, and quicker convergence, under fixed-pairs matching; the lone exception is Prisoners' Dilemma, where higher levels of C choices are associated with slower convergence to the unique Nash equilibrium (D,D).

### 3.2 Behavior dynamics

In Figure 2, we take a closer look at how behavior changes over time. This figure shows the relative frequency of each of the three types of outcome (both players choose C, both choose D, or exactly one chooses C) in each five-round block of each game, under each treatment. Arrows indicate the direction of motion of the time paths.

[Figure 2 about here]

As the figure shows, the dynamics of aggregate outcome frequencies vary substantially across games and between F and R treatments. Behavior begins near the point (1/4, 1/4) implied by uniform random play, but in most cases then moves in the direction of one of the pure-strategy pairs. In CG, both time paths move in the general direction of the Pareto efficient (C,C) Nash equilibrium, but the path for the R treatment ultimately gets closer, suggesting that some pairs in the F treatment become "stuck" at the Pareto inefficient (D,D) equilibrium.[5] In BoS, convergence toward the Nash equilibrium (C,D) and (D,C) pairs is faster and more uniform in the F treatment, while in PD, convergence toward the Pareto dominated Nash equilibrium (D,D) is faster in the R treatment.

In the three Stag Hunt games, qualitative differences between treatments are more pronounced. In SHH and SHM, time paths under random matching converge almost completely to the Pareto inefficient (D,D) outcome, while under fixed pairs, play converges roughly to equal frequencies of the (D,D) and the Pareto efficient (C,C) outcomes (suggesting heterogeneity across pairs); in both treatments, miscoordination on the (C,D) and (D,C) pairs has nearly died out by the end of the game. In SHL, the time path for the F treatment moves in the direction of (C,C), but a few pairs get stuck at (D,D) instead. The path for the R treatment shows some tendency toward roughly equal frequencies of (D,D) and (C,C) outcomes (that is, some *sessions* get stuck at the inefficient (D,D) outcome), but the frequency of miscoordination stays comparatively high.

By and large, therefore, Figure 2 confirms that the matching mechanism has not only a quantitative effect on the speed of convergence, but also a qualitative effect on the outcome to which the subjects converge. Miscoordination often persists even in late rounds—either due to incomplete convergence or to mixed-strategy play (either at the individual or population level)—but differences between F and R treatments are still apparent. In five of the six games, play converges to a better aggregate outcome (the three Stag Hunt games), more quickly to the same good outcome (Battle of the Sexes), or more slowly to the same bad outcome (Prisoners' Dilemma) under fixed pairs than under random matching. Only in one game (Coordination Game) does average behavior tend toward a worse outcome under fixed pairs. This happens because a nontrivial fraction of pairs get stuck at the lower-payoff equilibrium, as compared to random matching where all sessions converge to near-complete play of the higher-payoff equilibrium.

### 3.3 Parametric statistics

We next estimate a set of probit models—one for each game—with a C action choice as the dependent variable. Our primary explanatory variable is an indicator for the F treatment. We include the round number and its square, in order to control for changing behavior over time, as well as the number of subjects in the session, which in the R treatment influences how many times a subject can expect to face a given opponent (and thus potentially how past results affect choices). We also include all two- and three-way interaction terms among the F indicator, the time variables, and the session size (i.e., the number of subjects in the session). Additional controls are indicators for two of the three game orderings (M-L-H and L-H-M) and one of the two action orderings (C-D). Finally, we include the subject's previous-round own and opponent action choice, and their product; since each game is 2x2, these are sufficient to capture any dependence on the previous-round outcome. Each probit is estimated on the data from rounds 2-40 (not round 1, due to the previous-round variables) using Stata (v. 12), with a constant term and individual-subject random effects. Table 3 shows the results: estimated marginal effects (at variables' means) and standard errors for each variable, and the pseudo-$R^2$ for each model.

[Table 3 about here]

We see strong evidence of a treatment effect in the F-treatment marginal, which is significantly different from zero at the 5% level for five of the six games, and in each of these has the expected sign: negative in CG and positive in PD and the three SH games. Noting from earlier the lack of initial differences between the F and R treatments, the significance of the F-treatment variable implies differences in learning over time between these treatments. Figure 3 presents additional evidence for differences in learning, by showing additional marginal effects for the F-treatment indicator, for each individual round rather than at the mean round number as before. In the figure, point estimates are shown as dark circles, and line segments represent 95% confidence intervals. The effect of the F treatment on C choices clearly varies across games, but it starts out near zero in all games, and except for BoS, becomes significant by the second half of the session. In CG, the effect is negative, reflecting some pairs getting stuck on the (D,D) outcome, while in the other four games, the effect is positive; in all five of these games, the sign and significance of the effect persists throughout the 40 rounds. The size of the effect peaks in the first half of the session in PD, around the middle of the session in SHM, and near the end of the session in SHL; in the other three games it continues grows over time even at the end.

[Figure 3 about here]

To summarize, these results broadly replicate the patterns discussed in Section 3.2. Typically, outcomes are better, and learning is faster, in the F treatment than in the R treatment. Differences between the treatments are small in early rounds (before subjects have gotten feedback about the games' payoffs), but become more pronounced over time. In some cases, these differences grow over the course of a game; in others, they shrink in later rounds, as behavior in the R treatment "catches up" to that in the F treatment.

## 4. Discussion

In any strategic situation that is not so trivial that decision makers immediately figure out which actions to choose, it is important to be able to model the way their decision-making behavior adjusts over time (learning). In order to successfully model learning, we need to understand which aspects of the situation influence the way individuals learn. In this paper, we examine the one aspect, the matching mechanism: either *fixed-pairs matching* or *random matching*. Unlike most earlier investigations of the effects of the matching mechanism, subjects in our experiment play under *limited information*. They are never shown any game's payoff matrix, and while they receive information about their own payoffs in the end-of-round feedback, they never receive information about opponent payoffs. Besides ameliorating any effects of subjects' other-regarding preferences, our design serves to isolate the effect of the matching mechanism on learning, by limiting the effects of other factors that affect behavior, such as signaling or reputation building.[6]

In the experiment, we find sizable and systematic differences in behavior between fixed pairs and random matching. Outcomes are typically better under fixed pairs, as cooperative choices are more likely. However, the opposite can happen (as in our Coordination Game), when fixed pairs leads some pairs of subjects to become "stuck" on an inefficient equilibrium. We typically find faster convergence to pure-strategy play under fixed pairs, with the exception of Prisoners' Dilemma. Treatment effects are small to nonexistent in early rounds, but grow over time, and are not only visible in summary statistics, but confirmed by both non-parametric tests and parametric regressions. As noted in Section 2.1, the literature is fairly evenly split regarding whether the matching mechanism matters in limited-information settings, so given our results, it is fair to say that the preponderance of evidence is now in favor of the matching mechanism having an effect.

Our results can—to a fair degree—be explained by two distinct properties of fixed pairs relative to random matching. First, learning about opponent behavior is a simpler task when there is only one potential opponent (fixed-pairs matching) than when there are multiple potential opponents (random matching). This greater simplicity increases the likelihood of faster convergence to a Nash equilibrium under fixed pairs, as observed in most of the games. Second, groups of size two (fixed pairs) are more likely to show between-group heterogeneity in within-group average behavior than larger groups (random matching). When a game has multiple equilibria, the increased heterogeneity can lead to different groups converging to different action pairs, with the implication that some groups become stuck on inefficient equilibria. Indeed, of the games with multiple pure-strategy Nash equilibria, all five saw different pairs in the F treatment converging to different equilibria, but only two saw different sessions in the R treatment reaching different equilibria.

Our limited-information design was intended to minimize the effects of factors other than learning, in contrast to previous tests of matching mechanisms (see Section 2.1) that used complete-information designs. As a result, any differences between our results and those from complete-information experiments are likely to be due to these other factors. Some of our results (e.g., usually faster convergence to pure-strategy play under fixed pairs) have counterparts in complete-information experiments. However, our finding that the possibility of pairs becoming stuck on inefficient equilibria can lead to *worse* outcomes under fixed pairs than under random matching is generally not observed under complete information. While researchers have found that behavior in complete-information experiments can also be sensitive to early-round outcomes (Van Huyck et al., 1990), the usual consequence is that some subjects under fixed pairs use early rounds to signal cooperative actions (Clark and Sefton, 2001), improving outcomes relative to random-matching treatments, where incentives for such signaling are much weaker. Under limited payoff information, subjects cannot be confident that their opponents will successfully interpret signals of cooperative actions, making such signaling more difficult and less likely to be effective, leading to the possibility of worse outcomes under fixed pairs.

Though our main result involves a comparison of learning between two matching mechanisms, we note also that there was a general tendency under both mechanisms for play to move toward equilibrium. This convergence despite limited payoff information is consistent with similar results in the literature (Cox et al., 2001; Oechssler and Schipper, 2003; Shachat and Walker, 2004). We acknowledge, however, that our setting—with subjects playing symmetric 2x2 games with feedback including the opponent action and the own payoff—was especially conducive to subjects' gaining understanding of the strategic environment.

The most notable puzzle in our results comes from the Prisoners' Dilemma, where there was actually slower convergence to equilibrium under fixed pairs than under random matching. Figure 2 showed a persistent, non-negligible fraction of (C,C) outcomes, which in a complete-information setting would suggest some combination of social preferences or supergame behavior. However, our experiment was designed with limited information in order to avoid these possibilities, so it is unclear why there should be so many cooperative action choices.[7]

A few other remarks are warranted. First, we wish to encourage more work on this topic. We consider a study of the two most widely-used matching mechanisms in six fairly well-known games to be good progress toward understanding the effect of matching mechanism on learning but there are other matching mechanisms, alternative implementations of limited information, and countless other games that could be studied. Additionally, our study is idiosyncratic in some ways. With 6 games varied within-subject, we could not use all 6!=720 possible orderings; using only three leaves open the risk of order effects (i.e., behavior depending on what games were previously played, and differently under fixed pairs versus random matching). Future work might use fewer games per subject and all possible orderings, or only one game per subject, removing the need for variation of orderings entirely. Our study also had a high fixed payment to subjects relative to the variable payments. Since learning can be sensitive to stake sizes (e.g., Slonim and Roth, 1998), researchers may wish to use higher variable stakes to test the robustness of our conclusions. Finally, our use of symmetric games made it possible for subjects who figured out their own payoffs to make (correct) inferences about opponent payoffs; future work might focus on asymmetric games, so that guessing opponent payoffs would be much more difficult.

Second, our results have implications for the design of learning models. As discussed in the introduction, a successful model of learning ought to be able to predict the differences in play seen in our results. We would welcome further research that examines the ability of learning models to predict the differences between fixed pairs and random matching that we observed.[8]

Third, since the evolution of behavior over time is sensitive to the matching mechanism, we suggest that care should be taken whenever drawing conclusions based on data using only one matching mechanism. In cases where the only effect is on the speed of convergence, qualitative conclusions ought to be fairly robust. However, we've seen that the matching mechanism can have an effect on the outcome to which behavior converges, so that conclusions about equilibrium selection based on only one mechanism may be misleading. We are not arguing that all experiments should involve multiple treatments with varying matching mechanisms, as we recognize that experimental sessions are costly, and adding treatments with other matching mechanisms will generally imply a sacrifice of other potential design treatments. However, a possible compromise might be the use of additional pilot sessions involving different matching mechanisms, in order to assess the robustness of results to the mechanism used.

**References**


Ahn, T.K., Ostrom, E., Schmidt, D., Shupp, R., and Walker, J. (2001) “Cooperation in PD games: fear, greed, and history of play,” Public Choice 106, 137-155.

Andreoni, J., and Croson, R. (2008) “Partners versus strangers: random rematching in public goods experiments,” In C. Plott, and V.L. Smith (Eds.), Handbook of Experimental Economics Results, v. 1 New York: Elsevier, 776-783.

Camerer, C. (2003) Behavioral Game Theory: Experiments in Strategic Interaction, Princeton, NJ: Princeton University Press.

Camerer, C., and Ho, T.-H. (1999) “Experience-weighted attraction learning in normal form games,” Econometrica 67, 827-874.

Charness, G., and Garoupa, N. (2000) “Reputation, honesty, and efficiency with insider information: an experiment,” Journal of Economics and Management Strategy 9, 425-451.

Chen, Y. (2003) “An experimental study of serial and average cost pricing mechanisms,” Journal of Public Economics 87, 2305-2335.

Clark, K., and Sefton, M. (2001) “Repetition and signalling: experimental evidence from games with efficient equilibria,” Economics Letters 70, 357-362.

Danz, D. N., Fehr, D., and Kübler, D. (2012) “Information and beliefs in a repeated normal-form game,” Experimental Economics 15, 622-640.

Dekel, E., Fudenberg, D., and Levine, D.K. (2004) “Learning to play Bayesian games,” Games and Economic Behavior 46, 282-303.

Duffy, J., and Ochs, J. (2009) “Cooperative behavior and the frequency of social interaction,” Games and Economic Behavior 66, 785-812.

Erev, I., and Barron, G. (2005) “On adaptation, maximization, and reinforcement learning among cognitive strategies,” Psychological Review 112, 912-931.

Erev, I., and Roth, A.E. (1998), “Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria,” American Economic Review 88, 848-881.

Feltovich, N. (2000) “Reinforcement-based vs. beliefs-based learning in experimental asymmetric-information games,” Econometrica 68, 605-641.

Feltovich, N. (2005) “Critical values for the robust rank-order test,” Communications in Statistics— Simulation and Computation 34, 525-547.

Fischbacher, U. (2007) “z-Tree: Zurich toolbox for ready-made economic experiments,” Experimental Economics 10, 171-178.

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

Gerber, A. (2006) “Learning in and about games,” Working paper 234, Institute for Empirical Research in Economics, University of Zurich.

Hanaki, N., Ishikawa, R., and Akiyama, E. (2009) “Learning games,” Journal of Economic Dynamics and Control 33, 1739-1756.

Huck, S., Müller, W., and Normann, H.-T. (2001) “Stackelberg beats Cournot: on collusion and efficiency in experimental markets,” Economic Journal 111, 749-765.

Kagel, J.H., and Roth, A.E. (1995) Handbook of Experimental Economics, Princeton, NJ: Princeton University Press.

Kaneko, M., and Kline, J.J. (2008) “Inductive game theory: a basic scenario,” Journal of Mathematical Economics 44, 1332-1363.

McKelvey, R., and Palfrey, T. (2001) “Playing in the dark: information, learning, and coordination in repeated games,” working paper, California Institute of Technology.

Nicklisch, A. (2011) “Learning strategic environments: an experimental study of strategy formation and transfer,” Theory and Decision 71, 539-558.

Oechssler, J., and Schipper, B. (2003) “Can you guess the game you are playing?” Games and Economic Behavior 43, 137-152.

Rick, S., and Weber, R. (2010) “Meaningful learning and transfer of learning in games played repeatedly without feedback,” Games and Economic Behavior 68, 716-730.

Rieskamp, J., and Otto, P.E. (2006) “SSL: a theory of how people learn to select strategies,” Journal of Experimental Psychology. General 135, 207-236.

Roth, A.E., and 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.

Sarin, R. and Vahid, F. (2001) “Predicting how people play games: a simple dynamic model of choice,” Games and Economic Behavior 34, 104-122.

Schmidt, D., Shupp, R., Walker, J.M., and Ostrom, E. (2003) Playing safe in coordination games: the roles of risk dominance, payoff dominance, and history of play. Games and Economic Behavior 42, 281-299.

Siegel, S., and Castellan Jr., N.J. (1988) Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

Slonim, R. and Roth, A.E. (1998) Learning in high stakes ultimatum games: an experiment in the Slovak Republic. Econometrica 66, 569-596.

Van Huyck, J.B., Battalio, R.C., and Beil, R.O. (1990) Tacit coordination games, strategic uncertainty, and coordination failure. American Economic Review 80, 234-248.

Table 1: Characteristics of Nash equilibrium play

| Game | Strategy pair | Prob(C choice) | Game | Strategy pair | Prob(C choice) |
|---|---|---|---|---|---|
| | (1, 1) | 1.000 | | | |
| CG | (0, 0) | 0.000 | PD | (0, 0) | 0.000 |
| | (1/3, 1/3) | 0.333 | | | |
| | (1,0) | 0.500 | SHH, | (1, 1) | 1.000 |
| BoS | (0,1) | 0.500 | SHM, | (0, 0) | 0.000 |
| | (3/8, 3/8) | 0.375 | SHL | (2/3, 2/3) | 0.667 |

Table 2: Observed C choice frequencies from experiment

| Game | Rounds | F treatment | R treatment |
|---|---|---|---|
| | 1-5 | 0.533 | 0.530 |
| CG | 36-40 | 0.789 | 0.891 |
| | All | 0.684 | 0.750 |
| | 1-5 | 0.454 | 0.463 |
| BoS | 36-40 | 0.494** | 0.441 |
| | All | 0.466*** | 0.431 |
| | 1-5 | 0.454 | 0.424 |
| PD | 36-40 | 0.206* | 0.041 |
| | All | 0.290** | 0.126 |
| | 1-5 | 0.454 | 0.478 |
| SHH | 36-40 | 0.392*** | 0.085 |
| | All | 0.404*** | 0.203 |
| | 1-5 | 0.421 | 0.483 |
| SHM | 36-40 | 0.457** | 0.033 |
| | All | 0.446** | 0.125 |
| | 1-5 | 0.519 | 0.511 |
| SHL | 36-40 | 0.784 | 0.461 |
| | All | 0.746 | 0.499 |

Table 3: Probit estimated marginal effects and std. errors (dependent variable = C choice; N=9126)

| | CG | BoS | PD | SHH | SHM | SHL |
|---|---|---|---|---|---|---|
| F treatment | −0.063** | 0.028 | 0.100*** | 0.127*** | 0.240*** | 0.267*** |
| | (0.027) | (0.047) | (0.019) | (0.035) | (0.028) | (0.035) |
| Round | 0.007*** | 0.002*** | −0.004*** | −0.004*** | −0.004*** | 0.001 |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) |
| Session size | −0.003 | 0.000 | 0.003 | −0.003 | −0.001 | 0.003 |
| | (0.003) | (0.005) | (0.002) | (0.003) | (0.002) | (0.003) |
| MLH | −0.040 | −0.021 | −0.022 | −0.114*** | −0.064*** | −0.162*** |
| | (0.032) | (0.054) | (0.018) | (0.032) | (0.021) | (0.043) |
| LHM | −0.033 | 0.003 | 0.005 | −0.109*** | 0.050* | −0.041 |
| | (0.034) | (0.057) | (0.022) | (0.035) | (0.029) | (0.041) |
| CD | −0.113*** | −0.010 | 0.033** | 0.050 | −0.012 | −0.025 |
| | (0.025) | (0.043) | (0.016) | (0.031) | (0.020) | (0.031) |
| Previous C choice | 0.139*** | 0.281*** | 0.220*** | 0.304*** | 0.229*** | 0.281*** |
| | (0.013) | (0.015) | (0.020) | (0.022) | (0.022) | (0.017) |
| Opp. prev. C choice | 0.134*** | −0.313*** | 0.034*** | 0.242*** | 0.210*** | 0.362*** |
| | (0.013) | (0.015) | (0.011) | (0.021) | (0.022) | (0.017) |
| Pseudo-$R^2$ | 0.11 | 0.10 | 0.21 | 0.20 | 0.28 | 0.18 |

** (**,***): Coefficient is significantly different from zero at the 10% (5%, 1%) level.*

Figure 1: Games used in the experiment

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 2,2 | 0,0 |
| | D | 0,0 | 1,1 |

Coordination Game (CG)

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 0,0 | 3,5 |
| | D | 5,3 | 0,0 |

Battle of the Sexes (BoS)

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 7,7 | 1,8 |
| | D | 8,1 | 4,4 |

Prisoners' Dilemma (PD)

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 7,7 | 1,5 |
| | D | 5,1 | 5,5 |

Stag Hunt—high (SHH)

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 5,5 | –1,3 |
| | D | 3,–1 | 3,3 |

Stag Hunt—medium (SHM)

| | | Player 2 | |
|---|---|---|---|
| | | C | D |
| Player 1 | C | 1,1 | –5,–1 |
| | D | –1,–5 | –1,–1 |

Stag Hunt—low (SHL)

Figure 2: Time paths of experiment aggregate outcome frequencies (five–round blocks)

*Large (small) circles represent averages of rounds 1–5 (6–10, 11–15, etc.); arrows indicate time trend.*

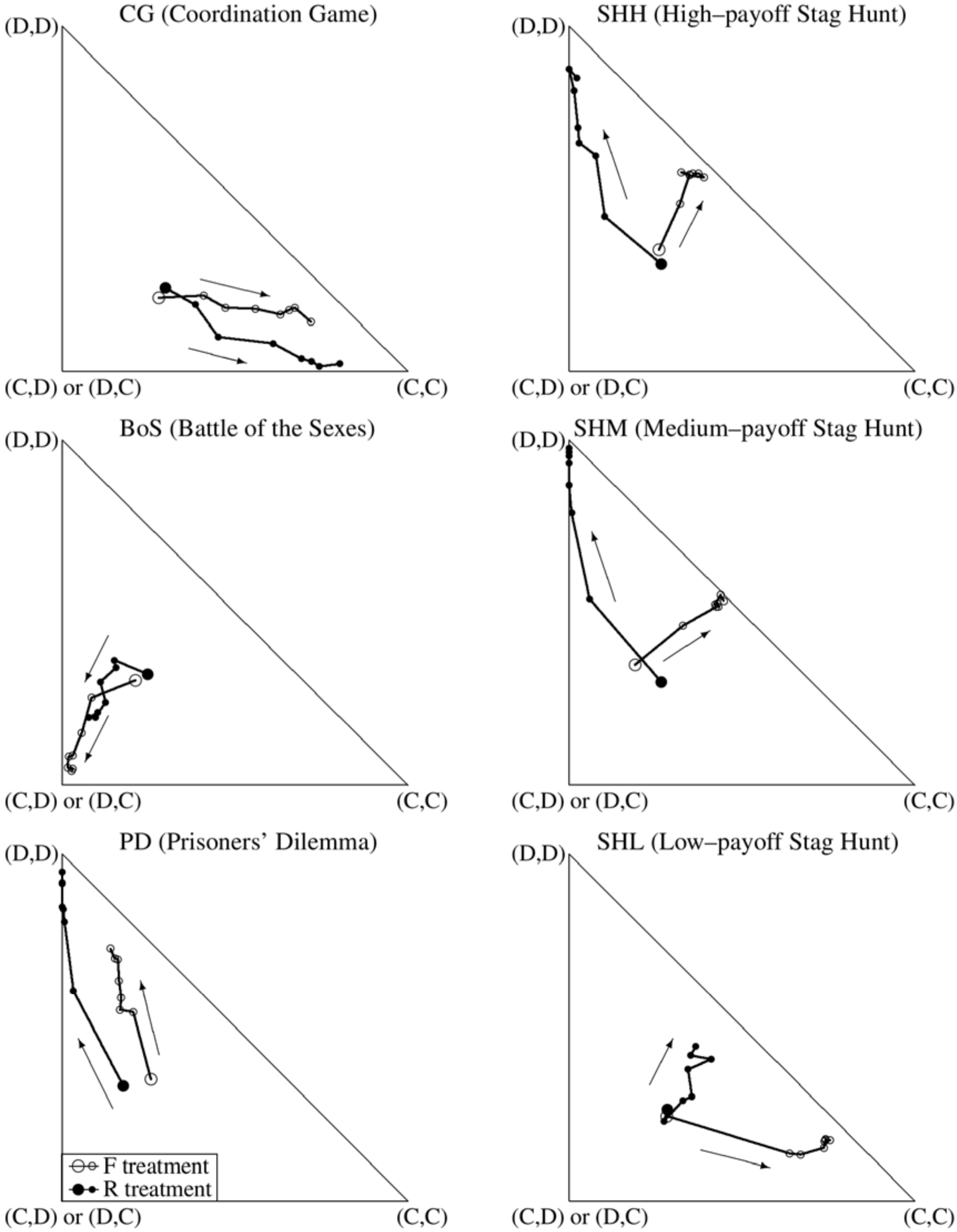
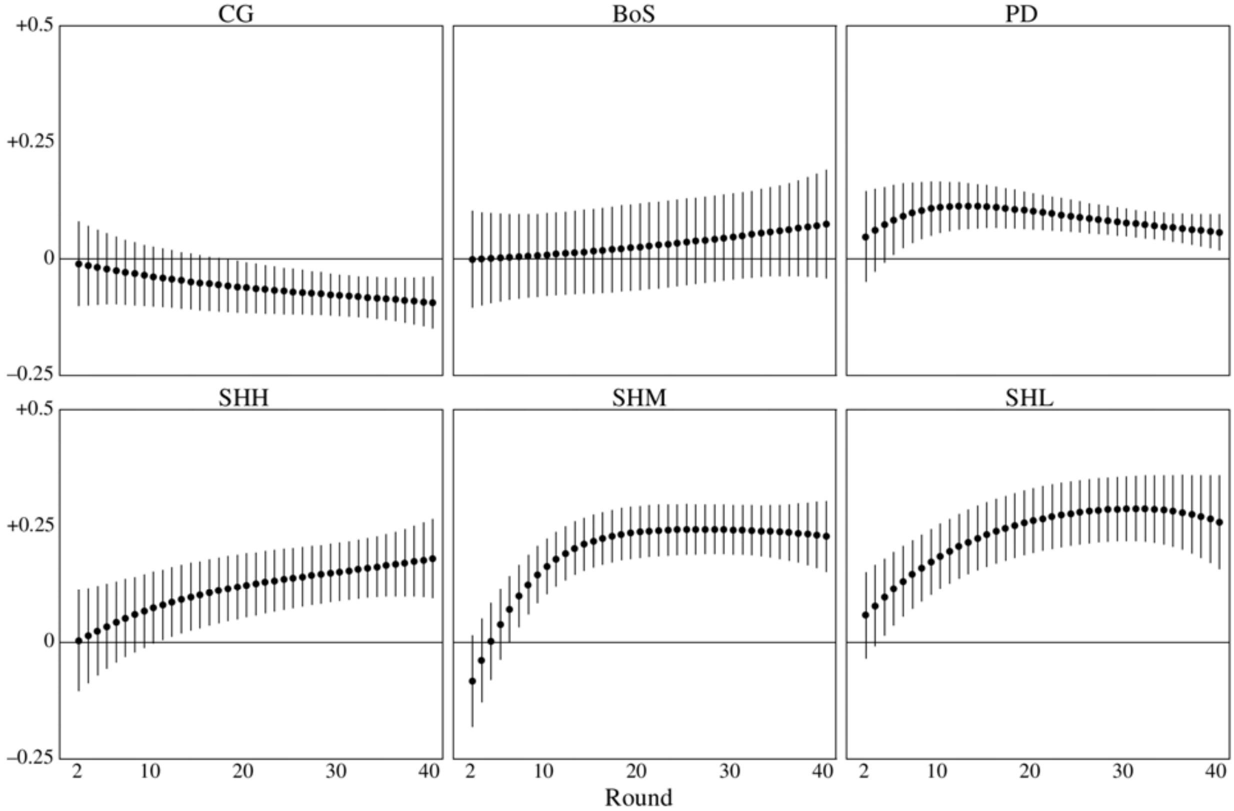


Figure 3: Estimated marginal effect of F (vs. R) treatment on C choice frequency in rounds 2–40, based on probit results *(circles represent point estimates; line segments represent 95% C.I.)*



[1] Corresponding author. Oda is supported in part by Grants-in-Aid of Scientific Research (#17310029 and #17700155) and by Japan's Ministry of Education, Culture, Sport, Science and Technology. Some of the research took place while Feltovich was at University of Houston and University of Aberdeen. We thank Atsushi Iwasaki for conducting the experimental sessions, and Ido Erev, an editor and two referees for helpful comments.

[2] Throughout this paper, we use the term "learning" to encompass any systematic change in behavior over time. This includes changes in play due to improved understanding of the structure of the game, as well as behavioral responses to changes in beliefs about how others play.

[3] They also had a "Partial Information" treatment where subjects were informed of their own portion of the payoff matrix, but they did not vary the matching mechanism there.

[4] On the other hand, Rick and Weber (2010) find that subjects learn iterated dominance even without feedback, and this understanding transfers across games.

[5] Of the 126 pairs of subjects in the F treatment, ten pairs play the (D,D) pair in each of the last 5 rounds, while two others do so in 4 of those rounds, suggesting that among a small but non-negligible number of pairs, behavior has converged to the (D,D) outcome.

[6] Because we did give subjects information about opponent action choices in the end-of-round feedback in our main experiment, we cannot completely rule out the possibility that some of the more sophisticated subjects were able to signal in early rounds via their action choices, and that these signals had an effect on opponent choices and coordination.

[7] One possibility is that some of the subjects assumed or inferred from feedback that the games were symmetric. In that case, they could then determine their opponents' payoffs once they had worked out their own part of the payoff matrix. If both subjects in a pair did so, then the chance of effects from social preferences or supergame behavior could increase substantially. One bit of evidence that suggests this might have happened is that of the 126 pairs in the F treatment of our original experiment, 18 played the (C,C) outcome in at least 9 of the final 10 rounds; that is, much of the (C,C) play was concentrated in a few pairs.

[8] An earlier version of this paper used a variant of Erev and Barron's (2005) RELACS model to characterize the data, with mixed success.

# Appendix A: instructions from the experiment

*Below is a annotated translation of the instructions given to subjects in this experiment. Copies of the actual instructions (in Japanese) are available from the corresponding author upon request.*

## 1 Introduction

Thank you for your participation in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions you might earn a considerable amount of money that will be paid to you in cash at the end of the session.

## 2 Sequence of Play in a Round

This experimental session consists of six different games. Each game consists of forty rounds. At the start of each game, you will be randomly assigned a player type, either "row player" or "column player." Your type will not change during the course of game. In each game of this experiment you will be randomly matched to a player of the opposite type. You will be matched with a different player in every round [every fortieth round in the F treatment] of a game. We will refer to the person you are paired with in a round as your "partner." Your score in each round will depend on your choice and the choice of your partner in that round. You will not know the identity of your partner in any round, even after the end of the session.

- At the beginning of each game, the computer program randomly matches each player to a partner.
- You and your partner play the game. Figure 4 is displayed on your screen. If you are a row player, you choose which row of the payoff table to play, R1 or R2. If you are a column player, you choose which column of the payoff table to play, C1 or C2.
- After all players have chosen actions, your action, your partner's action, and your payoff or score are displayed. Your score is determined by your action and the action of your partner according to the given payoff table.
- *[This part is replaced according to the treatment.]*
  - **R treatment:** Provided that the last round of the game has not been reached, a new round of the same game will then begin. You will be matched with a different partner in the new round.
  - **F treatment:** Provided that the last round of the game has not been reached, a new round of the same game will then begin. You will be matched with the same partner in the new round. Nevertheless, at every forty rounds, your partner and the payoff table will be changed.
- **Notice that you must not record any results of the games.** If the experimenter find you are recording them, you cannot continue your experiment. In that case, you will not be paid for this experiment.

[Figure 4 about here]

## 3 The payoff tables

The payoff table for each game you play will not be shown on your computer screen. However, let us explain the payoff table to support your decision-making. In every round of a game, both you and your partner have a choice between two possible actions. If you are designated as the row player, you must choose between actions R1 and R2. If you are designated as the column player, you must choose between actions C1 and C2. Your action, together with the action chosen by your partner, determines one of the four boxes in the payoff table. In each box, the first number represents your score and the second number represents your partner's score.

## 4 Payments

If you complete this experiment, the computer screen will reveal your score, the round that you got the score, and the payment you obtain. The round will be randomly chosen from all rounds you played. The payment will be calculated from your score as 200 yen for each point in that round. In addition, you will be paid a show-up fee of 3000 yen.

Are there any questions before we begin?

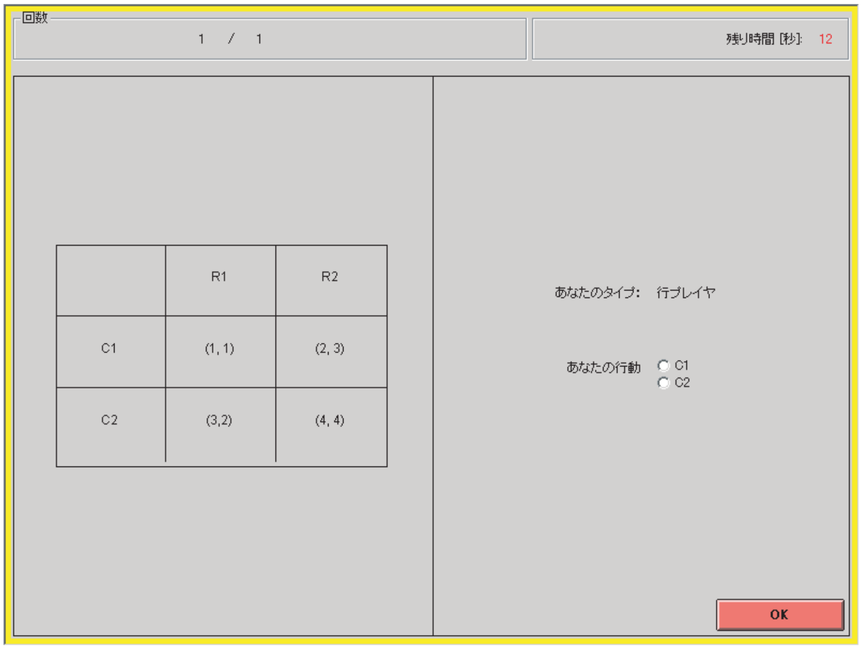



Figure 4: The screen when you are a column player. (The payoff table on the left side will not be shown)

# Appendix B: Session information

Table B1: Session information

| Session | Treatment | SH game ordering | Action ordering | Number of subjects | Aggregate C choice frequency | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | BoS | CG | PD | SHH | SHM | SHL |
| 1 | F | HML | CD | 28 | 0.468 | 0.586 | 0.541 | 0.464 | 0.587 | 0.811 |
| 2 | R | HML | CD | 20 | 0.451 | 0.646 | 0.089 | 0.345 | 0.144 | 0.816 |
| 3 | F | HML | DC | 22 | 0.473 | 0.864 | 0.253 | 0.457 | 0.374 | 0.740 |
| 4 | R | HML | DC | 20 | 0.415 | 0.798 | 0.149 | 0.264 | 0.135 | 0.470 |
| 5 | F | MLH | CD | 16 | 0.459 | 0.502 | 0.283 | 0.391 | 0.177 | 0.703 |
| 6 | F | MLH | DC | 6 | 0.467 | 0.912 | 0.012 | 0.350 | 0.058 | 0.662 |
| 7 | R | MLH | CD | 14 | 0.427 | 0.836 | 0.150 | 0.186 | 0.071 | 0.155 |
| 8 | R | MLH | DC | 10 | 0.425 | 0.672 | 0.152 | 0.088 | 0.135 | 0.190 |
| 9 | F | LHM | CD | 10 | 0.475 | 0.688 | 0.170 | 0.570 | 0.605 | 0.930 |
| 10 | F | LHM | DC | 18 | 0.474 | 0.704 | 0.204 | 0.339 | 0.636 | 0.642 |
| 11 | R | LHM | CD | 26 | 0.420 | 0.731 | 0.124 | 0.152 | 0.124 | 0.338 |
| 12 | R | LHM | DC | 18 | 0.450 | 0.814 | 0.111 | 0.126 | 0.131 | 0.851 |
| 13 | F | MLH | DC | 26 | 0.456 | 0.679 | 0.226 | 0.295 | 0.416 | 0.731 |

Treatments: F=fixed pairs, R=random matching.

SH orderings: second game played, fourth game played, sixth game played (CG always first, BoS always third, PD always fifth).

Action orderings: CD=C on top or left, DC=C on bottom or right.