First-Mover Advantage in Two-Sided Competitions: An Experimental Comparison of Role-Assignment Rules

Bradley J. Ruffle *Oscar Volij*

Department of Economics
Ben-Gurion University
Beer Sheva 84105
Israel

August 2012

Abstract:

Kingston (1976) and Anderson (1977) show that the probability that a given contestant wins a best-of-2k+1 series of asymmetric, zero-sum, binary-outcome games is, for a large class of assignment rules, independent of which contestant is assigned the advantageous role in each component game. We design a laboratory experiment to test this hypothesis for four simple role-assignment rules. Despite the fact that play does not uniformly conform to the equilibrium, our results show that the four assignment rules are observationally equivalent at the series level: the fraction of series won by a given contestant and all other series outcomes do not differ across the four rules.

Keywords: experimental economics, two-sided competitions, best-of series.

JEL Codes: C90, D02, L83.

Acknowledgments: This paper has benefitted from helpful conversations with Naomi Feldman, Guillaume Fréchette, Dan Friedman, Rosemarie Nagel, Oren Rigbi and numerous seminar and conference participants for valuable comments. Itai Carmon provided excellent research assistance. We are grateful to Ben-Gurion University for funding the experiments.

Contact Information:

Ruffle: Email: `bradley@bgu.ac.il`, Tel: +972-8-6472308.

Volij: Email: `ovolij@bgu.ac.il`, Tel: +972-8-6477346.

1 Introduction

Many sports and other two-sided competitions confer a strategic advantage to one side, typically the first mover. The serve in tennis and table tennis, the white pieces in chess and the home advantage in team sports like basketball, baseball and hockey are but a few examples. When two contestants compete in a best-of series, the question arises of how to assign them to the advantageous role in component games of the series. Consider, for instance, two contestants competing in a best-of-9 series in which the contestant in the role of Player 1 possesses an advantage in each component game. How does the rule that allocates contestants to roles in each game affect the outcome of the series?[1] Kingston (1976) and Anderson (1977) show that the probability that a given contestant wins the series is independent of the role-assignment rule for a large class of rules. In this paper, we report an experimental test of this equivalence theorem. In the experiment, paired contestants compete in a best-of-9 series of "Duel" under one of the following four theoretically equivalent role-assignment rules.

Alternating: contestants alternate in each game between the roles of Player 1 and Player 2;

5-4: one contestant plays the first 5 games in the role of Player 1 and any remaining games in the role of Player 2;

Winner: the winner of the current game assumes the role of Player 1 in the next game;

Loser: the loser of the current game assumes the role of Player 1 in the next game.

According to Kingston (1976) and Anderson (1977), the probability that the contestant who takes on the role of Player 1 in game 1 (to be referred to as the "leader") wins the series is the same for each of the above four assignment rules. More generally, they show that the probability that the leader wins a two-player series consisting of an odd number, $2k+1$, of identical, possibly asymmetric, zero-sum, binary-outcome games is independent of the rule that determines the identity of the contestant who plays in the role of Player 1 in each game. This result holds as long

[1] Nalebuff (1987) poses the related question of what constitutes a fair switching rule in table tennis when the two sides of the table are uneven and players switch sides only once.

as the rule does not assign either the leader to the role of Player 1 for more than $k+1$ games or the other contestant (to be referred to as the "follower") to the role of Player 1 for more than k games by the time the winner of the series is decided.[2] This clear-cut game-theoretic prediction rests on weak assumptions. In particular, the fact that the series consists of zero-sum games with only two outcomes implies that Kingston's theorem requires no special assumptions on players' risk preferences.[3]

We test whether these theoretically equivalent assignment rules are equivalent in the laboratory. Each subject plays eight best-of-nine series, each against a different opponent, under four different sets of game parameters. This setup provides us with a rich dataset to test Kingston's prediction, and its robustness over time and to the choice of game parameters. We also derive and test additional implications of the theory. For example, the probability that the winner of the first game also wins the series is predicted to be the same for the four role-assignment rules and independent of who won the first game. Furthermore, at the game level, we investigate for each role-assignment rule the extent to which individual play is consistent with equilibrium.

Several reasons suggest that behavior will differ significantly between the four assignment rules. To begin, their equivalence is premised on equilibrium play and subjects do not necessarily play according to equilibrium in a wide range of games (see Camerer 2003 for examples). Second, subjects may perceive *5-4* and *winner* as rules that favor the leader, while *alternating* and *loser* appear more even-handed. Psychological factors of this sort seem operative in a recent empirical literature that finds a non-negligible effect of the assignment rule on the outcome of the game. For instance, Magnus and Klaasen (1999) find an advantage to serving first in the first set of Wimbledon matches. Using data on professional soccer leagues and international tournaments, Apesteguia and Palacios-Huerta (2010) show that in penalty shootouts the probability that the team randomly chosen to shoot first wins is significantly higher than 1/2.[4] Feri et al. (2011) discover a second-

[2] To be clear about their respective contributions, Kingston (1976) demonstrates the equivalence between the *alternating* and *winner* rules. Anderson (1977) generalizes Kingston's result to show that any rule that meets the above condition is equivalent to *alternating*.

[3] Shachat and Wooders (2001) show the irrelevance of risk preferences for binary-outcome, repeated zero-sum games under general and weak conditions.

[4] On a different sample of soccer matches, Kocher et al. (2012) find that the first shooter's winning percentage is not significantly different from 1/2.

mover advantage in two-player free-throw shooting competitions in which the leader shoots five baskets one after the other and then the follower shoots his five baskets.

Notwithstanding, our results reveal strong support for the theory. The proportion of series won by the leader is similar for all role-assignment rules and similar to the theoretical point predictions. The same holds for the winner of game 1 whether leader or follower. This series-level equivalence across role-assignment rules is striking when contrasted with the observed differences in the quality of play across these rules at the game level: the frequency of equilibrium play is significantly higher in *winner* and *5-4* than in *alternating* and *loser*.

In the next section, we describe the series and its component games. We also demonstrate the theoretical equivalence between the four role-assignment rules. Section 3 details the experimental design and procedures. In Section 4, we present the hypotheses derived from the theory and the corresponding experimental results. Section 5 concludes.

2 The model

2.1 The stage game

The extensive-form version of the game "Duel" can be formalized as follows.[5] There are two players, each carrying a gun with a single bullet. The game tree has 20 sequential decision nodes. Player 1's decision nodes are the odd-numbered ones and those of Player 2 are the even-numbered ones. Formally, the players' sets of decision nodes are, respectively, $N_1 = \{1,3,\ldots,19\}$ and $N_2 = \{2,4,\ldots,20\}$. At each node except for the last one, the player whose turn it is to move decides whether to advance one step toward his opponent or to fire his gun. (In the last node, Player 2's only choice is to fire.) If he moves forward, the game continues to the next node. If, instead, player i fires at node $n \in N_i$, the probability of hitting his opponent is $p_i(n)$, for $i = 1,2$ and $n \in N_i$. The game ends as soon as one player fires his gun. This player becomes the shooter. If he hits his opponent, he wins and the other player loses. If he misses, he loses and his opponent is the victor. The probability functions p_i, $i = 1,2$, are assumed to be increasing in n, meaning that by delaying

[5] Binmore (2007) provides a lively analysis of Duel.

his shot, a player improves his chances of hitting, conditional on eventually shooting. By delaying his shot, however, he also allows his opponent the opportunity to fire first and thus end the game.

This game has a unique subgame-perfect equilibrium according to which contestant i plans to fire at every decision node $n \in N_i$ such that

$$p_i(n) + p_j(n+1) > 1, \qquad j \neq i, \tag{1}$$

and otherwise moves towards his opponent.[6] As a result, the equilibrium outcome involves a gun being fired at the first node n such that inequality (1) holds. In the experiment, we use the reduced normal-form representation of the game. Each contestant's action set consists of ten actions, each corresponding to each of his decision nodes. That is, $A_1 = \{1,3,\ldots,19\}$ and $A_2 = \{2,4,\ldots,20\}$. Each action represents the first node at which the contestant plans to fire his gun. Player 1's payoff function is

$$u_1(n_1,n_2) = \begin{cases} p_1(n_1) & \text{if } n_1 < n_2 \\ 1 - p_2(n_2) & \text{if } n_1 > n_2 \end{cases}$$

for $n_1 \in A_1$ and $n_2 \in A_2$. Player 2's payoff function is $u_2(n_1,n_2) = 1 - u_1(n_1,n_2)$. This game has a unique equilibrium, (n_1^*, n_2^*), corresponding to the unique subgame-perfect equilibrium of the extensive-form game described above. Since Duel is a zero-sum game,[7] by the minimax theorem (von Neumann 1928), there exists a value q^* such that the equilibrium action n_1^* of Player 1 guarantees that he wins with probability of at least q^*, and such that the equilibrium action n_2^* of Player 2 guarantees that he wins with probability of at least $1 - q^*$.

2.2 The series

A series consists of a sequence of multiple games in which two contestants, the *leader* and the *follower*, play $2k+1$ games of Duel with the winner of the series being determined by the contestant who wins $k+1$ games. We consider a best-of-9 series, namely, a contest in which the first

[6] We assume that $p_i(n) + p_j(n+1) \neq 1$ for every node. Otherwise, indifference between shooting and not shooting exists, thereby giving rise to an additional equilibrium.

[7] Strictly speaking, Duel is a constant-sum game. Since constant-sum and zero-sum games are strategically equivalent, we ignore this immaterial distinction and continue to refer to Duel as a zero-sum game.

contestant to win five games wins the series. The leader takes the role of Player 1 in the first game, while the follower assumes the role of Player 2. In the remaining games, the identity of Player 1 is determined by some specific rule. As previously mentioned, we consider four different rules. According to one rule, referred to as *alternating*, the leader plays in the role of Player 1 in the odd-numbered games and in the role of Player 2 in the even-numbered games. According to a second rule, referred to as *5-4*, the leader plays in the role of Player 1 in the first five games and in the role of Player 2 in the remaining four. A third rule, referred to as *winner*, assigns the winner of each game the role of Player 1 in the next game. Finally, *loser* is analogous to *winner*, except that from game 2 on, Player 1 is the contestant who lost in the previous game.

Note that the series is a finite zero-sum game. Therefore, by the minimax theorem, it has a value. More specifically, there exists a number p^*, such that there is a strategy for the leader that guarantees that he wins the series with probability of at least p^*, and there is a strategy for the follower that guarantees that he wins the series with probability of at least $1-p^*$.

A standard argument shows that playing the equilibrium action of Duel in each game constitutes an equilibrium of the series, independently of the four assignment rules under consideration.[8] To see this, let q^* be the equilibrium probability identified in the previous subsection that Player 1 wins the duel. Consider first the *alternating* rule. According to this rule, the leader will take on the role of Player 1 in five out of the nine component games. If he plays the equilibrium action in each of these five games, the probability that he wins exactly n of them is at least $B(5,n,q^*)$ where B stands for the binomial distribution. Similarly, by choosing his equilibrium action in each game he plays as Player 2, the leader can guarantee that the probability that he wins exactly m of these four games is at least $B(4,m,1-q^*)$. Therefore, if the leader plays his equilibrium action in each game, he will win the series with a probability of at least

$$P(win) = \sum_{n=0}^{5} \sum_{m=5-n}^{4} B(5,n,q^*)\, B(4,m,1-q^*). \tag{2}$$

Similarly, if the follower adopts the equilibrium action in each of the component games, he will

[8] Walker et al. (2011) provide a characterization of equilibrium strategies in general infinite-horizon, binary-outcome Markov games.

win the series with a probability of at least

$$P(lose) = \sum_{n=0}^{5} \sum_{m=5-n}^{4} B(5,n,1-q^*)\, B(4,m,q^*). \tag{3}$$

Routine calculations yield $P(win) + P(lose) = 1$, showing that the value of the series under the alternating rule, p^*, is $P(win)$. This value can therefore be attained by playing the equilibrium action in each component game.

The exact same argument applies to *5-4*, and more generally, to any rule according to which the leader is assigned the role of Player 1 in exactly five games (and the role of Player 2 in the four remaining games). Call these rules *balanced rules*. To see that this same argument extends to *winner* and *loser* as well, we can employ Anderson's (1977) ingenious reasoning. We refer to the repetition at which the winner of the series is determined as the "decisive duel". It can be seen that under both *winner* and *loser*, up until (and including) the decisive duel the leader has played as Player 1 at most five times, and the follower has played as Player 1 at most four times. Consider the following modification of the *winner* rule. The *modified winner* rule is like *winner* until the decisive duel. After the decisive duel, the roles are assigned so that the leader ends up playing as Player 1 exactly five times (and the follower exactly four times). By construction, the *modified winner* rule is a balanced rule. Thus, by the argument used above, p^* is the value of the series under the *modified winner* rule. Furthermore, it is clear that any two strategies, one for the *winner* rule and one for the *modified winner* rule, which coincide up to the decisive duel, yield the same probability of winning the series for the leader. Consequently, adopting the equilibrium action in each of the component games yields the same probability of winning the series under both the *winner* and the *modified winner* rules. Therefore, p^* is the value of the series under the *winner* rule as well. An analogous argument shows that the series also has the same value under the *loser* rule.

Kingston's theorem provides us with one clear testable implication, namely that the proportion of series won by the leader is independent of the role-assignment rule. But there are other implications as well. For instance, the probability that the winner of the first game ends up winning the series is also independent of the role-assignment rule, as well as of the identity of the contestant (leader or follower) who won the first game. These and other implications of equilibrium behavior will be tested in the next sections.

Prm. Table	Series		Player	Stage																			
	Sessions 1,2	Sessions 3,4		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
practice			1	.05		.10		.15		.20		.35		.55		.75		.85		.95		1	
			2		.06		.12		.18		.24		.30		.36		.42		.48		.54		.60
1	1, 5	4, 8	1	.04		.24		.42		.58		.72		.79		.85		.90		.94		.97	
			2		.05		.15		.25		.35		.45		.55		.65		.75		.85		.95
2	2, 6	3, 7	1	.09		.20		.31		.42		.53		.64		.75		.86		.97		1	
			2		.03		.07		.11		.15		.19		.23		.27		.31		.35		.39
3	3, 7	2, 6	1	.34		.48		.55		.68		.75		.83		.88		.93		.97		1	
			2		.02		.11		.25		.39		.53		.64		.73		.80		.85		.88
4	4, 8	1, 5	1	.02		.06		.14		.25		.51		.70		.74		.78		.82		.86	
			2		.02		.04		.08		.14		.20		.28		.36		.48		.60		.74

Table 1: Game parameterizations for the one practice and eight paid series of Duel. For each parameter table, each entry indicates the probability that the given player wins the game if he is the shooter at the given stage.

3 The experiment

3.1 Experimental design

To test Kingston's equivalence theorem, we design four experimental treatments that differ in the method of assignment to the advantageous role of Player 1. These treatments, discussed in Sections 1 and 2.2, will be referred to as *alternating*, *5-4*, *winner* and *loser*. We conduct four sessions of each treatment. In each session, pairs of subjects play eight best-of-nine series of Duel preceded by a practice series.

The parameters for these nine series are displayed in Table 1. Each entry shows the probability that the given player hits his opponent (and consequently wins the game) if he shoots at the corresponding stage. To illustrate, consider parameter table 1 (used in series 1 and 5 of sessions 1 and 2 and in series 4 and 8 of sessions 3 and 4). Suppose Player 1 plans to shoot at stage 5 and player 2 at stage 14. Player 1 becomes the shooter (because $n_1 < n_2$, in the notation of Section 2) and wins the game with probability .42 (alternatively, player 2 wins with probability .58).

In all sessions of all treatments, we employ the same set of game parameters. The parameters for the practice series are displayed in the first row of Table 1. For the eight paid series, we employ four distinct sets of game parameters. The choice of different parameters avoids basing our results on a single set of parameters and allow us to test the robustness of our results. Each parametrization appears twice, once within the first four series and again in the final four series. The ordering of these four parameter tables in sessions 1 and 2 is counterbalanced in sessions 3 and 4. Common to all of our chosen parameterizations is that they confer an advantage to the contestant in the role of Player 1. Namely, Player 1's equilibrium probability of winning an individual game exceeds .5 in all parameterizations. Moreover, our chosen parameter tables are such that Player 1's advantage is preserved as long as neither contestant deviates from the Nash equilibrium action by more than one stage. Also, even if both contestants choose randomly at which stage to fire, Player 1 maintains an advantage in each of the parameter tables.

These parameter tables differ by the identity of the shooter in equilibrium (Player 1 or Player 2), the stage in which the shooter shoots and the costliness (in terms of foregone probability) of deviating from the equilibrium action. The cost of deviating from equilibrium is high in two of the four parameter tables and low in the other two. Specifically, suppose the two players choose their equilibrium actions. If either player unilaterally deviates by one stage resulting in a change in the identity of the shooter, then the deviating player loses seven probability percentage points in parameter tables 1 and 3 (high cost) and two probability percentage points in tables 2 and 4 (low cost). Note that the higher the cost of deviation, the easier it should be for subjects to arrive at their equilibrium actions.

At this point, a comment is in order about our choice of game parameters. We have chosen parameters that provide the desired degree of difficulty for subjects in order to put forth an appropriately challenging test of the theory. If we chose a game with an easy equilibrium solution, subjects would play equilibrium in every game in all treatments. Consequently, play would trivially back Kingston's equivalence theorem. Instead, we have designed a game that many subjects may have difficulty arriving at the equilibrium solution. Indeed the stochastic nature of Duel admits two forms of misleading end-of-the-game feedback: a player who chooses the equilibrium action may lose the game and a player who deviates from equilibrium may win the game. At the same time, we expect some subjects to solve for (through iterative reasoning) or to intuit the equilibrium solution, while others may reach it through learning notwithstanding the misleading feedback. On the whole, we believe that the games we have designed strike an appropriate balance that gives both the null hypothesis and its alternative a fair chance to be rejected. The ultimate test of the suitability of our choice of parameters lies in the fraction of subjects who play equilibrium. A proportion not different from chance (i.e., 10%) would suggest that our game is too difficult for subjects, whereas almost everyone playing equilibrium in all games would raise suspicion that the theory would not withstand more challenging environments.

Within a series, the same pair of subjects plays Duel repeatedly until one of them wins five games. One pair member (termed the leader) is randomly assigned to the advantageous role of Player 1 in game 1. The treatment then determines the identity of Player 1 in all remaining games of the series. In subsequent series, the leadership is alternated from series to series such that each subject is the leader in exactly four of the eight paid series and in one of the two appearances of each parameter table.

Each subject faces a different opponent in each series (i.e., perfect strangers design). To implement this, we recruited groups of eighteen students and randomly divided them into two groups of nine. Group 1 students were leaders in the odd-numbered series and followers in the even-numbered series. Each student in group 1 played exactly one series against each of the students in group 2. In order to avoid any systematic ordering effect in the pairings,[9] we paired subjects with the help of a fixed but arbitrary solution to a Sudoku puzzle. Specifically, let A be the 9×9 matrix of the Sudoku solution with generic element a_{ij}. The rows of A represent the students of group 1, and the columns represent the students of group 2. The pairing is as follows: student i in group 1 plays against student j of group 2 in his a_{ij}th series. Since the entries of A are integer numbers between 1 and 9 such that each row contains one and only one of each digit, and similarly, each

[9] For example, we wish to avoid that contestant i's opponent in one series systematically plays against contestant j in the next series.

column contains one and only one of each of the nine digits, the above pairing is well-defined.

3.2 Experimental Procedures

All experiments were conducted in the Experimental Economics Laboratory at Ben-Gurion University using z-Tree (Fischbacher 2007). The treatment (i.e., role-assignment rule) was held constant throughout all series of a session. Four sessions were conducted for each treatment. The subject recruitment software limited participation to one session per subject. Eighteen subjects participated in each session, implying a total of 72 subjects per treatment and 288 subjects overall.

At the beginning of each session, printed instructions explaining the rules and the computer interface were handed out to subjects who were asked to read them carefully.[10] Then one of the experimenters read them aloud, after which the subjects answered a computerized comprehension quiz. One practice series was conducted for which the subjects received no payment followed by the eight paid series. Subjects received 10 NIS for each series they won plus a 30 NIS participation payment immediately after the session. With eight paid series played in pairs, the average subject could be expected to win four series for a total payment of 70 NIS.[11] The entire experiment, including the instruction phase and payment, lasted up to two hours and 15 minutes.

4 Results

4.1 Series-Level Results

We begin with an overview of series outcomes for each of the four experimental treatments. Since the results from the last four series do not differ dramatically from those based on all eight series – a testament to the difficulty of learning in this stochastic environment foreseen in the discussion in Section 3.1 – we use the complete dataset of eight series for all analyses. Table 1 displays the

[10] The instructions for the *alternating* treatment appear in the Appendix.
[11] At the time the experiments were conducted, $1 USD equalled approximately 3.5 NIS.

average length of a series and the distribution of final scores for each treatment. The first row of the table shows that series in *loser* lasted 7.81 games on average, the longest of any treatment followed closely by *alternating* at 7.6 games. Series in *winner* were resolved the quickest in 6.85 games. This ordering of treatments coincides precisely with the ordering of their theoretical expected lengths, which appears in the right-hand column of the first row for each treatment.

The remaining rows in the table display the distribution of final scores across treatments compared to the theoretical distribution. There are several noteworthy differences in final scores between treatments. Twenty-eight percent (81/288) of all series played under *winner* end in a 5-0 clean sweep compared to .003% (1/288) of all series in *loser*. These percentages are not out of line with those expected: 70.4 clean sweeps predicted in *winner* compared to only 2.6 in *loser*. At the same time, only 37% of all series in *winner* go to the eighth or decisive ninth game versus 54% in *5-4*, 57% in *alternating* and 65% in *loser*. χ^2-tests reveal that the distributions of final scores in *alternating* and *5-4* are not significantly different from the theoretically predicted distributions ($p = .12$ and $p = .59$, respectively), whereas the distributions in *winner* and *loser* are significantly different from their theoretical counterparts ($p = .01$ in both cases).[12]

Despite these differences between treatments, the following four results demonstrate that the treatments are virtually identical in the probability that a given contestant wins the series.

Hypothesis 1 (Kingston): The proportion of series won by the leader is the same for all treatments.

Result 1: The first row of Table 2 displays the fraction of series won by the leader over all series for each of the treatments. This fraction ranges from .545 (*winner*) to .580 (*alternating* and *5-4*).

[12] The reader may object to the use of the χ^2-test on the basis that each series is regarded as an independent observation. Specifically, even though each pair of contestants plays only one series together, every subject plays a total of eight series and play may be influenced by earlier series. Two retorts are possible. First, the theory assumes that individuals play equilibrium and consequently their decisions are independent of one another and across series. Second, alternative tests that treat the session as the unit of observation have less statistical power and are generally less likely to reject the equivalence of the four treatments. (See Fréchette (forthcoming), however, for an exception characterized by within-session variance that exceeds the variance of the session means.)

A χ^2-test of proportions reveals that the observed frequency with which the leader won the series does not differ significantly across treatments ($\chi^2(3) = 1.17$, $p = .76$).

In addition to testing the overall equivalence of the role-assignment rules, each set of game parameters affords a separate test.

Hypothesis 2: The proportion of series won by the leader is the same for all treatments in each of the parameter tables.

Result 2: The remaining rows of Table 2 display the fraction of series won by the leader separately for each of the parameter tables. We cannot reject the equivalence of the four treatments for any of the four parameter tables (p-values from χ^2-tests range from .34 to .96).

Hypotheses 1 and 2 follow directly from Kingston's result, which states that the probability that the leader wins the series is independent of the role-assignment rule. An analogous result holds regarding the winner of the first game. Concretely, the probability that the winner of the first game wins the series is independent of the treatment. Moreover, this probability is the same regardless of whether the leader or the follower won the first game of the series.[13] To see this, recall that the role-assigning methods *winner* and *loser* are equivalent to balanced rules (see Section 2.2). Therefore, it is sufficient that the statement holds for balanced rules. Consider a balanced rule and assume that contestant A wins the first game. In order for A to win the series, he must also win at least four of the eight remaining games. Since the role-assigning method is balanced, contestant A (whether leader or follower) will take on the role of Player 1 in exactly four of these remaining games. Therefore the probability that he wins the series is equal to the probability of winning at least four out of eight games, four of which he will play as Player 1. This probability ($\sum_{n=0}^{4} \sum_{m=n-4}^{4} B(4,n,q^*)B(4,m,1-q^*)$) is independent of whether A is the leader or the follower.[14] Hypotheses 3 and 4 address this extension.

[13] In other words, the leader's advantage in the series is restricted to game 1 of the series. In game 2, a contestant's probability of winning the series depends only on whether he won or lost game 1 and not on his role in game 1.

[14] This result does not generalize to games after the first one. For example, at the end of game 2, the probability that the contestant ahead in the series 2-0 goes on to win the series depends on whether the contestant is the leader or

Hypothesis 3 (Extension of Kingston): The proportion of series won by the winner of the first game is the same for all treatments.

Result 3: The first column of Table 3 shows that the proportion of series won by the winner of the first game ranges from 64.6% to 68.8% across the four treatments, with no significant differences between them ($\chi^2(3) = 1.24$, $p = .74$). Subsequent columns reveal that if the leader won the first game, the likelihood that he also won the series is approximately the same across treatments, varying between 65.4% and 68.5% ($\chi^2(3) = 0.79$, $p = .85$). Similarly, if the follower won game 1, the comparable range of percentages is from 63.1% to 73.8% with no significant difference between treatments ($\chi^2(3) = 2.99$, $p = .39$).

The next hypothesis claims that the above result holds even after conditioning on the role of the contestant who won game 1.

Hypothesis 4 (Extension of Kingston): The proportion of series won by the winner of the first game is independent of whether he is the leader or the follower.

Result 4: The row labeled "Overall" in Table 3 shows that, conditional on winning the first game, the chances of winning the series differ by less than a single percentage point for the leader (66.9%) and the follower (67.6%) ($\chi^2(1) = 0.06$, $p = .80$). Within each treatment (first four rows of Table 3), χ^2-tests of proportions show that if the leader won the first game, the likelihood that he went on to win the series does not differ significantly from the corresponding likelihood for the follower in any of the treatments (p-values are .68, .53, .17 and .70 for the respective treatments).

Thus far, we have conducted 14 statistical tests comparing the proportion of series won by a contestant across treatments. All 14 tests fail to reject the equivalence of the role-assignment rules at conventional significance levels. With between 72 and 288 observations in each cell for each of the tests performed, we would appear to have sufficient statistical power to reject the null. To show that this is indeed the case and to demonstrate additional support for the theory, we perform

follower and on the role-assignment rule.

these same tests across treatments on series outcomes not predicted to be the same. As footnote 14 indicates, conditional on the partial score at the end of game 2, the proportion of series won by the leader is expected to diverge across role-assignment rules.

Hypothesis 5: For each possible partial score at the beginning of game 3, the proportion of series won by the leader differs across treatments.

Result 5: For each of the four treatments, Table 4 shows the fraction of series won by the leader for each possible partial score at the beginning of game 3, namely, 2-0, 1-1 and 0-2 (where the digit before (after) the dash corresponds to the number of games won by the leader (follower)). For the partial score 2-0, the first row of the table indicates that the leader went on to win 73.3%, 78.1% and 80.0% of the series in *5-4*, *alternating* and *winner*, respectively. In *loser*, this win percentage vaults to 93.3%. Consequently and for the first time up to this point, the win frequencies are significantly different from one another ($\chi^2(3) = 10.5$, $p = .02$).[15] For the partial score 1-1 (second row of Table 4), a χ^2-test also rejects the equality of the win frequencies ($p = .01$), owing largely to the relatively high percentage of series (60.9%) won by the leader in *alternating* (between 12 and 19 percentage points higher than the other three treatments). Only after a partial score of 0-2 (third row of Table 4) does the χ^2-test fail to reject the equality of the frequency with which the leader won the series ($p = .38$). Results 1 – 4 show that the theory correctly predicts the equivalence of the role-assignment rules. The point of Result 5 is to show that when the theory predicts that the assignment rules are not equivalent, indeed they are not.

The first five results compare either the likelihood of a given contestant winning the series across treatments or, in the case of Result 4, the likelihood of different contestants winning the series within a treatment. If ours was a field experiment, the series winner would be the sole basis for testing Kingston's equivalence result since the underlying probabilities of winning a game and

[15] Contrast this highly significant difference with the parallel result after game 1 (reported in Result 3 and seen in the middle column of Table 3), according to which the proportion of series won by the leader after winning the first game (1-0) is statistically indistinguishable across treatments ($p = .85$).

the overall series would be unobservable. No further test of Kingston's result would be possible. We therefore would conclude that our field experiment unequivocally affirms the theory. However, one advantage of our laboratory experiment – and lab experiments more generally – is that the underlying game parameters are observable and generate a wealth of additional predictions related to the hypothesized equivalence of the role-allocation rules.

Results 6 and 7 present additional series-level analyses, which continue to support Kingston's theorem.

Hypothesis 6: The proportion of series won by the leader equals the theoretical probability.

Result 6: The rows of Table 2 display the theoretical probability and corresponding fraction of series won by the leader for each set of game parameters and aggregated over all game parameters for each of the treatments. In the aggregate, the realized fractions of series won by the leader range from .545 to .580, depending on the treatment. None of these fractions differs significantly from the theoretical prediction of .562 (Binomial test p-values from .55 to .77). Looking at the separate parameter tables, the observed fractions resemble the respective theoretical predictions in most cases and indeed only one of the discrepancies is significant at the 10% level or less – *winner* in parameter table 2, $p = .02$. With 16 tests performed, one rejected null hypothesis is in line with the number to be expected, namely, 0.8 expected rejections at 5% and 1.6 at 10%.

The theoretical probability that the leader wins a series is based on the assumption that both the leader and the follower play their equilibrium actions in every game. In the next subsection, we explore the extent to which this strong assumption holds. In the meantime, we will evaluate whether our data can reject alternative behavioral assumptions. Its failure to do so would suggest that hypotheses other than equilibrium play are also consistent with observed behavior, thereby weakening the support for equilibrium play as the likely explanation for our findings. Each alternative behavioral assumption that we will consider involves a small deviation from equilibrium play. For example, suppose the follower always wants to be the shooter. To achieve this, whenever he is not the shooter in equilibrium (i.e., Player 1 in parameter tables 1 and 2, Player 2 in parameter tables 3 and 4), he deviates by firing a single stage earlier. Under this behavioral assumption, the resultant probability that the leader wins the series aggregated over all parameter tables increases by six percentage points to .621. Comparing this theoretical probability to the observed fraction of series won by the leader, we can reject the equality between the two for two of the four treatments (*winner* and *loser*) ($p < .02$ in both cases), while we cannot quite reject the equality between the two in *alternating* and *5-4* ($p = .16$ in both cases).

A second alternative to equilibrium play is that the follower never wants to be the shooter. Accordingly, whenever the equilibrium dictates that he is the shooter, he delays his shot by one stage. As a result of this deviation, the leader's probability of winning the series increases to .620. Again, we reject the equality between this probability and the observed fraction of series won by the leader for two of the four treatments. Two additional alternatives to equilibrium play are also rejected by the Binomial tests. If the leader always wants to shoot first, his probability of winning the series drops to .503, whereas if he never wishes to shoot first, the corresponding probability falls to .502. We can reject at the 10% level of significance the equality of these respective probabilities and the observed fractions of series won by the leader for three of the four treatments in both cases. In sum, even a single-stage deviation in only about half of the games of the series yields significant inconsistencies with our data. *A fortiori* for more substantial deviations.

Hypothesis 7: The probability that the winner of the first game goes on to win the series equals the theoretical probability and is independent of the player's role.

Result 7: The first column of Table 3 shows that the proportion of series won by the contestant who won the first game of the series ranges from 64.6% (*loser*) to 68.8% (*alternating*). Binomial tests reveal that none of these percentages differs significantly from the theoretical prediction of 65.3% (Binomial test *p*-values range from .24 to .80). Although similar to the theoretical predictions, these overall treatment percentages may hide opposite tendencies between the leader and the follower that, on average, cancel out one other. This turns out not to be the case. The "Leader" and "Follower" columns in Table 3 suggest that each of the probabilities is similar to the theoretical probability of .653. In fact, none of the percentages for the leader (Binomial test p-values from .37 to 1) or for the follower (Binomial test p-values from .13 to .91) differ significantly from .653.

Until now, our focus has been on comparing treatments to one another and to the theoretical point predictions at the series level. Our results reveal strong support for the theory: the proportion of series won by the leader is similar for all treatments and similar to the theoretical prediction. And the same holds for the winner of game 1 whether leader or follower. The remainder of this section examines play at the game level where the predictive power of the theory reveals its first cracks.

4.2 Game-Level Results

The last column of the first row in Table 5 indicates that when aggregated across all games in all treatments, 56.1% of decisions correspond to the equilibrium. Consonant with our goal of choosing a game that is neither too easy nor too difficult for subjects, this percentage lies smack in the middle of the two extremes of random choice (10%) and full equilibrium play (100%). Moreover, most deviations are a single stage away from the equilibrium choice. In fact, play within one stage of the equilibrium accounts for 88% of decisions overall. The average absolute deviation from equilibrium (i.e., the absolute value of the difference between the chosen stage and the equilibrium stage) is 0.63 stages. Furthermore, in 41.3% of the games, the shot is fired at the equilibrium stage. In 35% of the games, both players chose their equilibrium actions. In the remainder of this subsection, we explore whether role-allocation rules differ in their frequency of equilibrium play.

Hypothesis 8: The frequency of equilibrium play is the same in all treatments.

Result 8: Table 5 highlights that, according to several distinct measures, play is better in *winner* and *5-4* than in *loser* and *alternating*. To begin, the percentage of games in which a contestant chose the equilibrium action is highest in *winner* (62.6%), followed closely by *5-4* (61.3%) and lowest in *loser* (52.2%) and *alternating* (49.4%).[16] In addition, the magnitude of the average absolute deviation from equilibrium is smallest in *winner* and *5-4*.

Results 1–7 all regard play in an individual series as the unit of observation. For comparability, we compute the frequency with which paired contestants choose the equilibrium action in a given series, consisting of between five and nine games (between 10 and 18 choices for the pair). Resembling closely the mean subject-game frequencies reported above, mean series-level frequencies of equilibrium play are 61.7% in *winner*, 61.2% in *5-4*, dropping off to 52.3% in *loser* and 49.4% in *alternating*. The non-parametric Wilcoxon-Mann-Whitney test reveals that the distributions of series-level frequencies are not significantly different in *winner* and *5-4* ($p = .75$) nor in *loser* and *alternating* ($p = .21$); however, any other two treatments are significantly different from one another (all $p < .01$).

Despite these treatment-level differences in equilibrium play, the equilibrium action is without exception the modal choice in each treatment and in each parameter table as well as all combinations thereof. For each player (1 and 2), the equilibrium action is not only the optimal choice against an opponent's equilibrium action, it also turns out to be an optimal choice against the opponent's observed distribution of actions in the population for each parameter table overall as well as for each treatment separately. If we compare behavior within one stage of equilibrium, 90.4% and 89.2% of decisions in *5-4* and *winner*, respectively, correspond to this more lenient measure of equilibrium play, compared to 86.7% and 85.9% in *loser* and *alternating*. Because the equilibrium stages vary widely across parameter tables, these findings provide strong evidence that subjects do not play according to simple behavioral rules, such as "always fire in the middle stage."

Figure 1 provides further evidence that subjects play Duel sensibly: even their deviations from equilibrium adhere to some rationale. The figure plots the cumulative distributions of choices

[16] This same ordering holds when we rank treatments by the percentage of games in which: i) the leader chose his equilibrium action; ii) the follower chose his equilibrium action; iii) the shot was fired at the equilibrium stage; and iv) both contestants chose their equilibrium action.

expressed as deviations from the equilibrium action. Three distinct distributions are displayed: (i) the overall distribution of deviations (solid line); (ii) the distribution of deviations given that in the previous game of the same series the opponent chose to shoot late (i.e., after the equilibrium stage) (dashed line); and (iii) the distribution of deviations given that in the previous game of the same series the opponent chose to shoot at least two stages after the equilibrium stage (dotted line). Distribution (i) highlights graphically the above observation that about 90% of contestants' choices are within a single stage of the equilibrium. What is more, comparing distributions (ii) and (iii) with (i) reveals that contestants' choices are responsive to their opponents' lagged choices. If the opponent fired late in the previous game, the contestant tends to delay his shot in the current game – and the contestant's delay is even greater if the opponent fired at least two stages late. In fact, contestants' reactions to their opponents' delayed shot are sufficiently strong that the three distributions are ordered according to first-order stochastic dominance: (iii) dominates (ii) which dominates (i). If a contestant believes that his opponent will again fire late as in the previous game, then firing late is a rational response.[17]

We turn now to regression analyses to explain observed deviations from equilibrium. We estimate a linear probability model with random effects. The baseline model is as follows,

$$y_{igr} = \alpha_0 + \alpha_1 \mathit{5\text{-}4} + \alpha_2 \mathit{winner} + \alpha_3 \mathit{loser} + \beta x + u_i + \varepsilon_{igr}, \tag{4}$$

where the indices i, g and r represent the subject, game and series, respectively. The dependent variable y is equal to 1 if individual i in game g of series r chose the equilibrium action, and 0 otherwise. The independent variables *5-4*, *winner* and *loser* are binary indicators equal to 1 if the subject played in the corresponding treatment, and 0 otherwise. The vector x represents variables related to the game, series and contestant's role, all of which are discussed below. Finally, u_i is the subject-specific random effect, while ε represents the idiosyncratic error term. Standard errors

[17] The centipede game bears some resemblance to our "Duel" game in that each contestant wishes to move one stage before his opponent (as long as the move is not before the equilibrium stage) and given the contestant moves first his payoff increases monotonically in the stage that he moves. Similar to our findings, Nagel and Tang (1998) show that subjects in a repeated centipede game respond to their opponent's decision to move after them in a given round by (weakly) delaying their move in the next round.

are clustered by subject, taking into account the correlation in the error terms over the games and series within a subject. Table 6 presents the results.[18]

Regression (1) displays the marginal effects from three of the four treatments. The constant of .494 reflects the mean percentage of equilibrium play in the omitted treatment *alternating*; in *loser* this fraction is not significantly different from that in *alternating* ($p = .45$), whereas both *winner* and *5-4* reveal significantly higher frequencies of equilibrium play (13 and 12 percentage points higher, respectively) than *alternating* ($p < .01$ in both cases). A t-test of coefficients shows that *winner* and *5-4* are not significantly different from one another ($p = .76$). Thus, this and subsequent regressions confirm the above results from non-parametric tests. There appear to be two distinguishable groups of treatments in terms of frequency of equilibrium play: a relatively low-frequency group consisting of *alternating* and *loser*, and a high-frequency group consisting of *5-4* and *winner*.

One might conjecture that the likelihood of equilibrium play depends on whether the contestant is the leader in the series or player 1 in the game. Regression (2) shows that neither of these variables significantly affects the likelihood of equilibrium play. Moreover, the coefficients and significance levels of the treatment dummies remain unchanged when these controls are included.

Some features of the parameter tables might be thought to affect the likelihood of equilibrium play. For example, a higher opportunity cost of a one-stage deviation from equilibrium might induce fewer deviations from equilibrium. The coefficient of .039 ($p < .01$) in regression (3) indicates that moving from a low-cost to a high-cost parameter table reduces the frequency of deviation from equilibrium by four percentage points. Whether the equilibrium of the game dictates that Player 1 or Player 2 is supposed to be the shooter does not significantly affect the frequency of equilibrium play. Again the treatment effects remain robust in magnitude and significance to the inclusion of these controls.

[18] If instead of the linear probability model we estimate Probit regressions, the significance and non-significance of all coefficients in all of the reported regressions remain unchanged, which is not surprising given that almost all of our regressors are binary variables (Angrist and Pischke 2010). We report the former for ease of interpretation.

Not all games in a given series are equally important. Some games are pivotal, while the outcomes of others do not substantially affect a contestant's chances of winning the series. Morris (1977) proposes to measure the importance of a given game in a best-of-k series as the difference between the probability of a given contestant winning the series conditional on winning the game and the probability of the same contestant winning the series conditional on losing the game. Formally, let $P(s)$ be the probability that contestant A wins the series given that the series' partial score is s. After the game is played, there are two possible partial scores: the partial score that results if A wins the game, denoted s_w, and the partial score that results if A loses the game, denoted s_ℓ. The importance of the game with a partial score s is given by $P(s_w) - P(s_\ell)$. Note that since the probability that contestant B wins the series given any partial score s is $1 - P(s)$, the importance of the game is independent of the identity of the contestant ($P(s_w) - P(s_\ell) = 1 - P(s_\ell) - (1 - P(s_w))$).

To convey the meaning of the importance of the game, let us use the following analogy to betting in poker. Suppose winning the series is worth 1. Each contestant possesses an endowment equal to his current probability of winning the series given the partial score s. In particular, $P(s)$ represents the endowment of contestant A. Correspondingly, $1 - P(s)$ is contestant B's endowment. Each contestant places a wager on the current game such that if he loses, he will be left with the resulting probability of winning the series. Specifically, contestant A bets $P(s) - P(s_\ell)$, while contestant B stakes $P(s_w) - P(s)$. The winner of the game collects the sum of these wagers, which is exactly the importance of the game. In this sense, the importance of the game captures what is really at stake in the game.

Figure 2 provides a concrete illustration of this importance-of-the-game measure for each possible partial score based on the *5-4* treatment and parameter table 2. The figure highlights a number of features of this measure. First, when the series is tied 4-4, the ninth game becomes a winner-take-all game and therefore has an importance of 1. At the other extreme, when the partial score is 0-4, the fifth game has an importance close to 0. The reason is that if the leader loses the game, he loses the series; but even if he wins the game, his likelihood of winning the series is close to 0 because he needs to win the next four games, all as Player 2.

Regression (3) shows that the likelihood of equilibrium play increases with the importance of the game. The coefficient of 0.103 in (3) suggests that the transition from a game with importance 0.35 to the decisive game with importance 1 (e.g., in *5-4*, parameter table 2, the transition from game 8 with the leader behind 3-4 to game 9) increases the probability of equilibrium play by 6.7 percentage points. In addition, the ordering of treatments by frequency of equilibrium play is preserved and the significance or lack thereof of each treatment dummy remains unchanged with the inclusion of these variables.

The significance of both the high-cost and importance-of-the-game variables show that play improves with an increase in monetary incentives. Because the importance of the game tends to increase over the course of the series (see Figure 2), it could be that the positive association of this variable with the frequency of equilibrium play masks a learning effect: contestants' understanding of the game improves during the series resulting in better choices. Indeed, Figure 3 reveals that the overall fraction of choices that correspond to the equilibrium action tends to increase over the course of the series in each of the four treatments, especially from game 1 to game 2. To distinguish between the effects of game importance and learning, we include measures of both in regression (4).

Within a series, the likelihood of playing the equilibrium action increases by one percentage point from one game to the next ($p < .01$). From series to series, the proportion of equilibrium play rises by two percentage points ($p < .01$). On the other hand, the point estimate of $-.008$ for the importance-of-the-game variable is not significantly different from 0 ($p = .48$).[19] These findings imply that the source of the observed improved play as the series progresses is learning rather than improved performance from higher game stakes.

[19] The Spearman correlation coefficient of .33 between the game number in the series and the importance-of-the-game measure suggests that multicollinearity is not a concern. We also ran this same regression specification separately for each treatment. The series and game variables continue to be highly significant in each treatment. The importance-of-the-game variable is not significantly different from zero in three of the treatments ($p > .5$ in all three cases) and only marginally significant ($p = .10$) in *winner*.

The central insight from this subsection is that the observed quality of play varies significantly across role-allocation rules. The frequency of equilibrium play is highest in *5-4* and *winner* and lowest in *alternating* and *loser*. These differences between allocation rules would appear to belie the rules' similarities for series outcomes. The next regression reconciles these seemingly disparate findings.

Regression (5) includes indicator variables for the four treatments, each interacted with a dummy variable equal to 1 if the subject played in the role of follower for series *r*. The three treatment indicators from previous regressions are also present. Their coefficients are now to be interpreted as the difference in the leader's frequency of equilibrium play in the specified treatment from that in *alternating* (given by the constant). Applied to (5), the highly significant coefficients of .12 and .13 on *5-4* and *winner*, respectively, reveal that the leader in these treatments chooses the equilibrium action 12 and 13 percentage points more often than the leader in *alternating*. More importantly, the coefficients on the treatment-follower interaction terms are all tiny – the largest among them being .007 – and not significantly different from zero.[20] Simply stated, within each treatment, the leader and the follower each select the equilibrium action with equal frequency. Had the leader played equilibrium with a significantly different frequency than the follower in one or more, but not all, of the treatments, then we would expect the fraction of series won by the leader to differ across treatments, which subsection 4.1 shows is not the case.

Random-effects regressions parallel to those in Table 6 appear in Table 7 with the dependent variable y_{igr} in equation (4) being replaced by a binary indicator equal to 1 if in game g of series r subject i chose the equilibrium action or within one stage of it. This more inclusive definition of equilibrium play renders the treatments more similar to one another. In fact, the only two treatments that are even weakly significantly different from one another are *5-4* and *alternating* ($p = .09$ in each of (6) – (10) in Table 7). No other treatments can be rejected as being similar at conventional levels of significance. Regression (7) shows that Player 1 is significantly more

[20] This identical result (not included in the table) is obtained if we include some or all of the controls from the previous regressions.

likely to play within one stage of the equilibrium action than Player 2. None of the game-related variables including the importance of the game significantly affects the likelihood of play within one stage of the equilibrium according to (8). Similar to regression (4), (9) indicates the importance of learning: play improves from series to series and from game to game within a series, whereas the importance of the game continues not to differ significantly from zero. Regression (10) shows that the leader and the follower choose within one stage of the equilibrium action with frequencies that do not differ significantly from one another in three of the four treatments. In *5-4*, however, the follower plays within one stage of his equilibrium action two percentage points less often than the leader. These results in (10) are robust to the inclusion of controls from previous regressions.

5 Conclusions

Elimination series in team sports often follow a best-of format. To allocate the home advantage over the entire series as equally as possible, the two teams typically alternate in the role of the home team. Kingston (1976) and Anderson (1977) present a striking theorem that shows that whether the home advantage is alternated or allocated according to some other rule doesn't matter: all role-assignment rules belonging to a large class of rules yield the same probability that a given contestant wins the series. This equivalence holds under general conditions with almost no restrictions on contestants' preferences.

We design a laboratory experiment consisting of four dissimilar but theoretically equivalent assignment rules. Our results reveal strong support for the theorem at the series level. The proportion of series won by the leader is similar for all four assignment rules and similar to the theoretical point predictions. The same is true for the proportion of series won by the winner of game 1 whether leader or follower. This series-level equivalence holds despite significant differences in the frequency of departure from equilibrium play across assignment rules.

Suboptimal play in zero-sum games may arise from poor choice or poor execution. By poor choice we mean a player's deliberate choice of a non-equilibrium action. By contrast, poor execution refers to a situation in which a player intends to implement the equilibrium action, but unintentionally chooses another one. To illustrate the distinction, the penalty shooter in soccer may be fully cognizant of his optimal shooting strategy, but nerves may get the best of him in its implementation, sending the ball sailing over the crossbar (i.e., poor execution). Alternatively, for white to open a chess game with "pawn to a3" is an example of poor choice. In our experiment, the implementation of a contestant's choice is straightforward: simply type the stage number. Hence, we conclude that suboptimal play in our setting likely follows from poor choice. It would be interesting to extend this study of role-assignment rules to settings in which only poor execution is operative.

References

[1] Anderson, C. L. "Note on the Advantage of First Serve," *Journal of Combinatorial Theory* (A) 23, 363 (1977).

[2] Angrist, Joshua D. and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press (2009).

[3] Apesteguia, Jose and Ignacio Palacios-Huetra. "Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment," *American Economic Review* 100:5, 2548–2564 (2010).

[4] Binmore, Kenneth. *Playing for Real*, Oxford University Press (2007).

[5] Camerer, Colin. *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press (2003).

[6] Fischbacher, Urs. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics* 10:2, 171–178 (2007).

[7] Fréchette, Guillaume R. "Session-Effects in the Laboratory," *Experimental Economics* (forthcoming).

[8] Kingston, J.G. "Comparison of Scoring Systems in Two-Sided Competitions." *Journal of Combinatorial Theory* (A) 20, 357–363 (1976).

[9] Kocher, Martin G., Marc V. Lenz and Matthias Sutter. "Psychological Pressure in Competitive Environments: New Evidence from Randomized Natural Experiments," *Management Science* (forthcoming).

[10] Magnus, Jan R. and Frand J. G. M. Klaassen. "On the Advantage of Serving First in a Tennis Set: Four Years at Wimbledon," *The Statistician* 48, 247–256 (1999).

[11] Morris, C. "The Most Important Points in Tennis," in *Optimal Strategies in Sports*, S. P. Kadany and R. E. Machol (eds.), Amsterdam: North-Holland, 131–140 (1977).

[12] Nagel, Rosemarie and Fang Fang Tang. "Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning," *Journal of Mathematical Psychology* 42, 356–384 (1998).

[13] Nalebuff, Barry. "Puzzles," *Journal of Economic Perspectives* 1:1, 157–163 (1987).

[14] Shachat, Jason and John Wooders. "On The Irrelevance of Risk Attitudes in Repeated Two-Outcome Games," *Games and Economic Behavior* 34, 342–363 (2001).

[15] von Neumann, John. "Zur theorie der Gesellschaftsspiele," *Mathematische Annalen* 100, 295–300 (1928).

[16] Walker, Mark, John Wooders and Rabah Amir. "Equilibrium Play in Matches: Binary Markov Games," *Games and Economic Behavior* 71, 487-502 (2011).

Appendix: Participant Instructions

Introduction

This is a decision-making experiment. Funds for this experiment have been provided by various research foundations. Take time to read carefully the instructions. A good understanding of the instructions and well thought out decisions in the experiment can earn you a considerable amount of money. All earnings from the experiment will be paid to you in cash at the end of the experiment.

The Duel

In this experiment, you will play 8 matches, each one against a different opponent. Each match will be played as a best 5 out of 9 games, meaning that the first player to win 5 games wins the match. Your earnings will be determined by the number of matches you end up winning.

An example of a component game of a match appears in the table below.

Player	Stage																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	.05		.10		.15		.20		.35		.55		.75		.85		.95		1	
2		.06		.12		.18		.24		.30		.36		.42		.48		.54		.60

Note: Each entry in the table reveals the probability with which the given player wins the game if he shoots at the indicated stage and shoots before his opponent. His opponent wins with the complementary probability.

To understand this game, it will be useful to think of a duel between two shooters walking toward one another. Each shooter (player) has a gun with only one bullet and must choose when to fire (i.e. use his bullet) with the objective of hitting his opponent. Player 1 can shoot or advance toward his opponent at the odd-numbered stages only, while Player 2 can shoot or advance at the even-numbered stages only.

Both players decide simultaneously at which stage they intend to shoot. But in fact only one player actually gets to shoot: the player who decides to shoot at the earlier stage becomes the shooter. For instance, if Player 1 intends to shoot at stage 3 and Player 2 intends to shoot at stage 4, then Player 1 becomes the shooter.

The shooter hits his opponent, and consequently wins the game, with the probability indicated in the table and loses the game with the complementary probability. (We assume that if the shooter misses his opponent, he loses the game because he has no more bullets and his opponent can walk up to him and hit him with certainty.)

Suppose, for example, Player 1 intends to shoot in Stage 5 and Player 2 intends to shoot in Stage 6, then Player 1 becomes the shooter, winning the game with probability 0.15 and losing the game with probability .85. Alternatively, suppose Player 1 again intends to shoot in Stage 5, but Player 2 decides to shoot in Stage 4. In this case, Player 2 is the shooter and wins the game with probability .12 and loses with probability .88.

Note that the longer one waits before shooting, the higher is the probability of hitting. However, by not shooting at a given stage, a player gives his opponent an opportunity to shoot first in the next stage.

Method of Determining Player 1 in Each Game

One of you will be randomly selected to begin game 1 of the match as Player 1 with the other player assigned to the role of Player 2. In game 2 of the match, the roles will be reversed. Players will continue to alternate between the roles of Player 1 and Player 2 until the match is over. That is, one of you will be in the role of Player 1 in all odd-numbered games (and in the role of Player 2 in all even-numbered games); while the other will be in the role of Player 1 in all even-numbered games (and in the role of Player 2 in all odd-numbered games).

Matches

In total, you will play 8 matches for real preceded by one practice match. The probabilities for the practice match are given in the above table. For the 8 real matches, 4 different sets of probabilities will be used. More precisely, you will play 2 matches with each of the four sets of probabilities. In one of the two matches, you will assume the role of Player 1 in game 1. In the other of the two matches, you will assume the role of Player 2 in game 1. In which of the matches you will begin as Player 1 and in which you will begin as Player 2 will be determined randomly.

Overall, each player will begin as Player 1 in game 1 in 4 of the 8 matches. You will be shown the table of probabilities at the beginning of the each match.

In each of the 9 matches (1 practice and 8 real) you will face a different opponent. This opponent is someone against whom you will not have played in any previous match (including the practice match) and against whom you will not play in any future match. In each match, your opponent will be randomly determined from among those participants in the room against whom you have not already played.

Payments

After completing all 9 matches, you will be asked to complete a short questionnaire after which you will be paid your earnings from the experiment in cash. Everyone will receive a 30 NIS payment for having participated in the experiment. In addition, you will earn 10 NIS for each match you win (excluding the practice match). Note that the winner's payment is 10 NIS regardless of whether the final score of the match is 5-4, 5-0 or any score in between. You earn nothing (0 NIS) for matches you lose.

If at any stage you have any questions about the instructions, please raise your hand and a monitor will come to assist you. Before beginning the experiment, everyone will answer a brief quiz to ensure that they have understood the rules of the experiment.

Thank you for your participation.

Table 1 - Realized and predicted series lengths and final scores by treatment

	Alternating		5-4		Winner		Loser	
	Realized	Expected	Realized	Expected	Realized	Expected	Realized	Expected
Ave. Series Length	7.60 (1.24)	7.68	7.41 (1.46)	7.33	6.85 (1.55)	6.86	7.81 (1.02)	8.09
5-0	16	12.6	49	50.7	81	70.4	1	2.6
5-1	49	36.4	32	36.4	56	55.5	39	17.1
5-2	58	68.9	52	49.7	44	56.6	60	52.5
5-3	77	82.2	62	69.0	38	54.7	101	96.6
5-4	88	87.9	93	82.2	69	50.8	87	119.3
Total Obs.	288	288	288	288	288	288	288	288
χ^2 test statistic	7.31		2.82		15.98		38.86	
p-value	0.12		0.59		0.00		0.00	

The first row indicates the mean series length (standard deviation in parentheses) and the expected series length for each treatment. Subsequent rows display the distribution of final scores by treatment and a χ^2 test evaluating whether the observed and expected distributions of final scores differ significantly from one another.

Table 2 - Frequency that Leader won the series

Parameter Table	Theoretical	Alternating	5-4	Winner	Loser	$\chi^2(3)$ p-value
Overall	0.562	0.580	0.580	0.545	0.552	0.76
1	0.568	0.542	0.542	0.514	0.500	0.95
2	0.543	0.611	0.597	0.681**	0.597	0.69
3	0.581	0.583	0.597	0.528	0.611	0.76
4	0.553	0.583	0.583	0.458	0.500	0.34

Theoretical and observed frequencies that the leader won the series, by treatment and by parameter table.
** indicates observed frequency is significantly different from theoretical probability at 5% level (two-sided Binomial test).

Table 3 - Frequency of winning the series conditional on having won the first game

Treatment	Overall mean	Overall obs.	Leader mean	Leader obs.	Follower mean	Follower obs.	$\chi^2(1)$ p-value between Leader and Follower
Alternating	0.688	288	0.681	213	0.707	75	.68
5-4	0.674	288	0.685	197	0.648	91	.53
Winner	0.677	288	0.654	208	0.738	80	.17
Loser	0.646	288	0.654	185	0.631	103	.70
Overall	0.671	1152	0.669	803	0.676	349	.80
Theoretical prediction	0.653		0.653		0.653		
$\chi^2(3)$ p-value between treatments	.74		.85		.39		

The first column displays for each treatment the observed proportion of winning the series given that the contestant won the first game. In subsequent columns, these proportions are displayed separately for the leader and follower. The observations are the number of series in which the indicated contestant won the first game. The last column displays the p-values from χ^2 tests of the equality of the frequency with which the leader and follower won the series given each respectively won the first game, as predicted by the theory.

Table 4 - Frequency that Leader won series conditional on partial score after game 2

Partial Score at End of Game 2	Alternating mean	Alternating obs.	5-4 mean	5-4 obs.	Winner mean	Winner obs.	Loser mean	Loser obs.	$\chi^2(3)$ p-value
2-0	0.781	73	0.733	146	0.800	130	0.933	60	.02
1-1	0.609	169	0.471	104	0.416	101	0.493	201	.01
0-2	0.152	46	0.289	38	0.193	57	0.148	27	.38
Overall	0.580	288	0.580	288	0.545	288	0.552	288	

For each treatment and possible partial score at the beginning of game 3, the entries display the fraction of series won by the leader. The last column reports p-values from χ^2 tests of equality of these fractions.

Table 5 - Measures of equilibrium play

Equilibrium Play Measure	Alternating Mean (s.d.)	5-4 Mean (s.d.)	Winner Mean (s.d.)	Loser Mean (s.d.)	Overall Mean (s.d.)
% play eq'm	0.494 (0.500)	0.613 (0.487)	0.626 (0.484)	0.522 (0.500)	0.561 (0.496)
% play eq'm + 1 deviation	0.859 (0.348)	0.904 (0.295)	0.892 (0.310)	0.867 (0.339)	0.880 (0.325)
mean absolute deviation	0.726 (0.942)	0.533 (0.829)	0.557 (0.926)	0.695 (0.957)	0.631 (0.919)
% eq'm shot	0.337 (0.473)	0.468 (0.499)	0.490 (0.500)	0.368 (0.482)	0.413 (0.492)
% eq'm outcome	0.277 (0.448)	0.413 (0.492)	0.421 (0.494)	0.298 (0.458)	0.350 (0.477)

The first three rows concern individual play. % eq'm shot indicates the percentage of games in which the observed shot corresponds to the equilibrium stage. % eq'm outcome refers to the percentage of games in which both contestants chose their equilibrium actions.

Table 6
Regressions on the frequency of equilibrium play

Variable	(1)	(2)	(3)	(4)	(5)
5-4	0.119 ***	0.121 ***	0.116 ***	0.122 ***	0.122 ***
	(0.039)	(0.039)	(0.039)	(0.039)	(0.043)
winner	0.132 ***	0.13 ***	0.125 ***	0.135 ***	0.132 ***
	(0.040)	(0.040)	(0.040)	(0.040)	(0.043)
loser	0.029	0.031	0.032	0.03	0.034
	(0.038)	(0.038)	(0.038)	(0.038)	(0.041)
leader		-0.003			
		(0.009)			
player1		-0.002			
		(0.010)			
high-cost deviation			0.039 ***		
			(0.013)		
player1 eq'm shooter			-0.011		
			(0.009)		
game importance			0.103 ***	-0.008	
			(0.009)	(0.024)	
series				0.022 ***	
				(0.002)	
game				0.012 ***	
				(0.002)	
alternating*follower					0.007
					(0.020)
5-4*follower					0.005
					(0.017)
winner*follower					0.002
					(0.017)
loser*follower					0.000
					(0.029)
constant	0.494	0.494	0.444	0.444	0.488
	(0.026)	(0.026)	(0.027)	(0.027)	(0.029)
adj. R^2	.013	.013	.015	.027	.013
obs.	17092	17092	17092	17092	17092

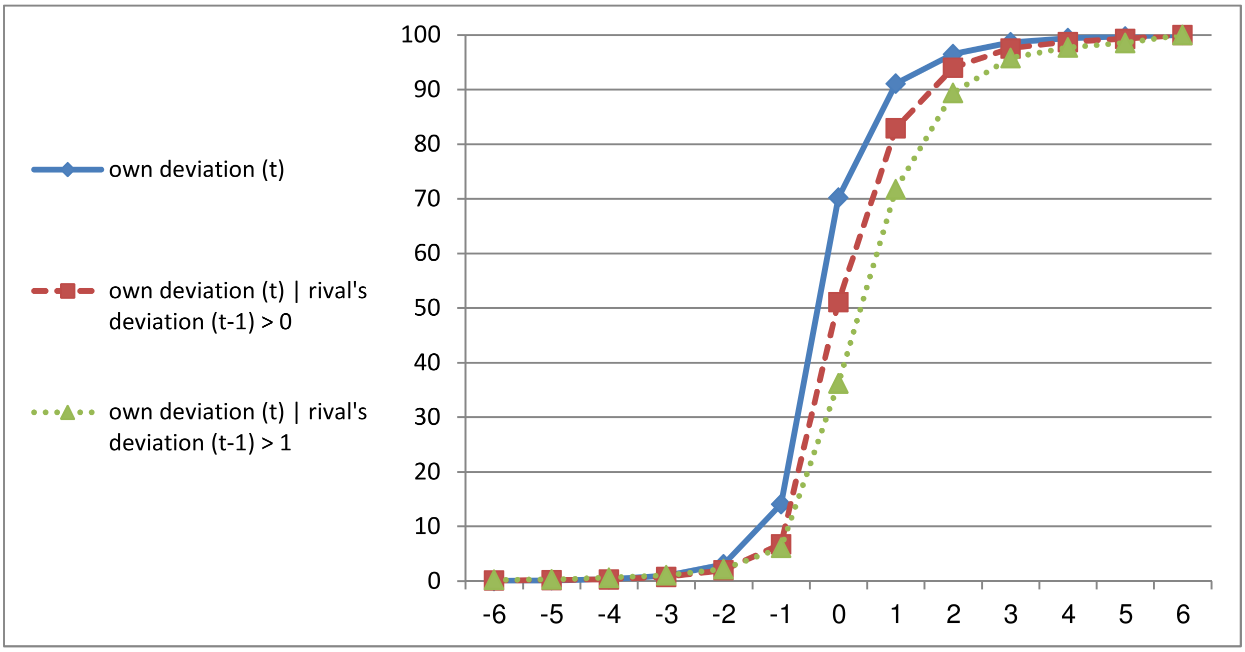
Notes: 1. Linear probability model with random effects.
2. Dependent variable: binary variable for whether subject *i* in game *g* of series *r* played the equilibrium action.
3. Independent variables: indicators for *5-4*, *winner* and *loser* treatments; indicators for whether subject *i* was the leader (*leader*) or player 1 (*player 1*); indicators for whether the parameter table involves a high cost of deviation from equilibrium, for whether player 1 is the equilibrium shooter; a measure of the importance of the game (*game importance*); the series and game numbers; interaction variables between treatment dummy and dummy for whether subject *i* was the follower.
4. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 7
Regressions on the frequency of equilibrium play or within one stage of it

Variable	(6)		(7)		(8)		(9)		(10)	
5-4	0.045	*	0.045	*	0.044	*	0.045	*	0.058	**
	(0.026)		(0.026)		(0.026)		(0.026)		(0.027)	
winner	0.033		0.033		0.032		0.034		0.040	
	(0.027)		(0.027)		(0.027)		(0.027)		(0.029)	
loser	0.008		0.008		0.008		0.008		0.019	
	(0.027)		(0.027)		(0.027)		(0.027)		(0.028)	
leader			0.000							
			(0.006)							
player1			0.048	***						
			(0.006)							
high-cost deviation					-0.004					
					(0.008)					
player1 eq'm shooter					0.001					
					(0.006)					
game importance					0.005		0.001			
					(0.006)		(0.016)			
series							0.011	***		
							(0.002)			
game							0.003	**		
							(0.001)			
alternating*follower									0.007	
									(0.016)	
5-4*follower									-0.020	**
									(0.009)	
winner*follower									0.009	
									(0.011)	
loser*follower									0.015	
									(0.014)	
constant	0.859		0.836		0.854		0.857		0.488	
	(0.020)		(0.021)		(0.020)		(0.020)		(0.029)	
adj. R^2	0.003		0.008		0.003		0.01		0.013	
obs.	17092		17092		17092		17092		17092	

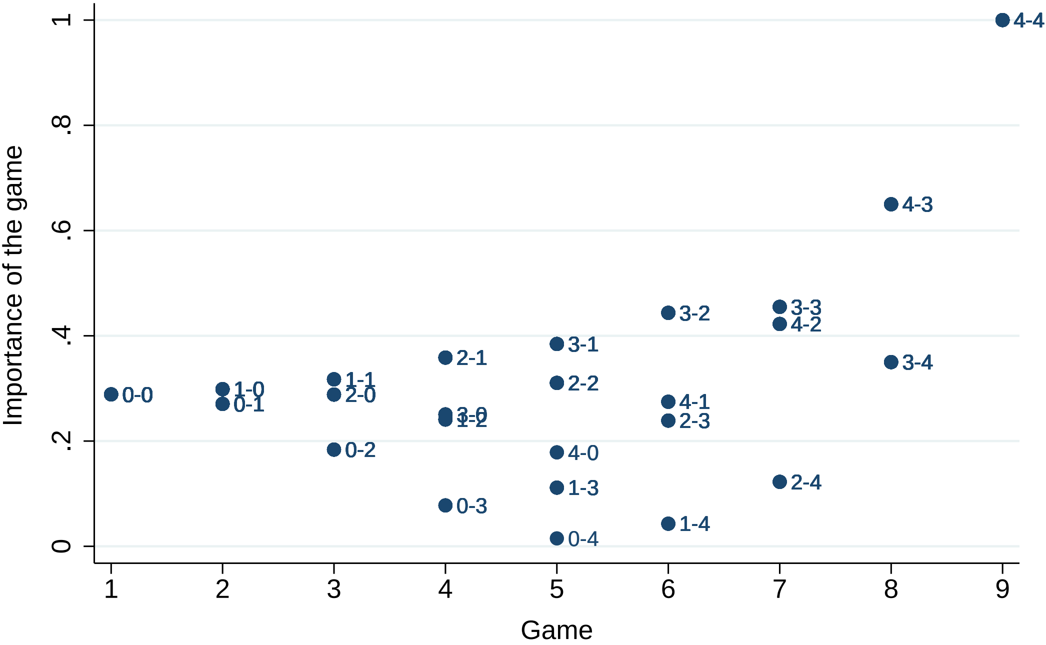
Notes: 1. Dependent variable: binary variable for whether subject *i* in game *g* of series *r* played the equilibrium action or within one stage of the equilibrium.
2. See Table 6 for further explanations.

Figure 1 - Cumulative distributions of deviations from equilibrium



Cumulative distributions of the contestant's deviation from the equilibrium stage displayed for all games, for games in which the opponent shot late in the previous game, and for games in which the opponent shot at least two stages late in the previous game.

Figure 2 - Importance of the game by partial score



Note: based on 5-4 treatment, parameter table 2

Figure 3 - Fraction of equilibrium choices by game for each treament