LCERPA

Laurier Centre for Economic Research & Policy Analysis

LCERPA Working Paper No. 2014-16

September 2014

First-Mover Advantage in Best-Of-Series: An Experimental Comparison of Role-Assignment Rules

Bradley J. Ruffle, Department of Economics, Wilfrid Laurier University
Oscar Volij, Department of Economics, Ben-Gurion University

First-Mover Advantage in Best-Of Series: An Experimental Comparison of Role-Assignment Rules

Bradley J. Ruffle
Department of Economics
Wilfrid Laurier University
Waterloo, ON, N2L 3C5
Canada
bruffle@wlu.ca

Oscar Volij
Department of Economics
Ben-Gurion University
Beer Sheva 84105
Israel
ovolij@bgu.ac.il

September 2014

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 significant differences in the frequency of equilibrium play across the four assignment rules, our results show that the four 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 rules.

Keywords: experimental economics, two-sided competitions, best-of series, asymmetric game, psychological pressure.

JEL Codes: C90, D02, L83.

Acknowledgments: This paper has benefited from helpful conversations with J. Atsu Amegashie, Naomi Feldman, Guillaume Fréchette, Dan Friedman, Andrew Healy, Logan McLeod, Rosemarie Nagel, Oren Rigbi, Justin Smith, Ro'i Zultan, numerous seminar and conference participants, two anonymous referees and an associate editor at this journal for valuable comments. Itai Carmon provided excellent research assistance. We are grateful to Ben-Gurion University for funding the experiments.

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 the component games of the series. Consider, for instance, 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) answer this question by showing 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 under one of four theoretically equivalent role-assignment rules, to be referred to as *alternating*, *5-4*, *winner* and *loser*. The *alternating* rule requires that contestants alternate between the roles of Player 1 and Player 2 in every game, whereas the *5-4* rule places one contestant in the role of Player 1 for the first five games before switching to the role of Player 2 for any remaining games. The contestant who wins the current game assumes the role of Player 1 in the next game according to *winner*, whereas *loser* awards the loser of the current game with the advantaged role 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 as the rule does not assign the leader to the role of Player 1 for more than $k+1$ games nor the

[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.

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 the Kingston-Anderson theorem requires no special assumptions about 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 and Anderson'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 may 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* may 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 chosen to shoot first wins is significantly higher than 1/2.[4] Feri et al. (2011) discover a second-mover advantage in two-player free-throw shooting competitions in which the leader shoots five baskets

[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.

one after the other and then the follower shoots his five baskets. Finally, the more asymmetric the component games, the more divergent partial scores are likely to be after, say, game 4, in series played according to *5-4* than in those played according to more balanced rules like *alternating*. If the partial score affects behavior, then play may differ across different assignment rules. In this paper we test the Kingston-Anderson equivalence theorem in an environment in which observed winning probabilities may vary across treatments due to frequent departures from equilibrium play which result from divergent partial scores or other perceived differences between the role-assignment rules.

Three theoretical contributions reveal that unequal partial scores indeed bear consequences for the effort contestants elect to invest in subsequent games. Harris and Vickers (1987) propose two models of multi-stage races. At every stage of the race, two players exert effort as a result of which one of the players wins the stage. In one model, the winner of the race is the first one who wins a predetermined number of stages. In the other, the objective is to achieve a predetermined score difference. Harris and Vickers (1987) show that under certain conditions the equilibrium dictates that the player who is ahead exert more effort than the one who is behind. Klumpp and Polborn (2006) examine the dynamics of candidate behavior as a function of the partial score in the U.S. primaries. If effort is costly, they find that the outcome of the first election induces the winner to invest more effort in subsequent primaries, while the loser exerts less effort. Konrad and Kovenock (2009) study the equilibrium of a multi-battle contest in which players participate in a sequence of all-pay auctions. As long as the component auctions award a small prize to the winner, they show that very unequal partial scores induce higher expected effort than close scores.

Several experimental and empirical studies have tested the predictions of the above models.[5] Zizzo (2002) and Mago et al. (2012) report experiments on some versions of the Harris and Vickers (1987) models. While the first study does not find evidence that the leader spends more effort than the follower, the second one does. Irfanoglu et al. (2010) experimentally test the predictions of Klumpp and Polborn (2006) and find substantial evidence of the New Hampshire effect, namely, that a best-of-three contest is more likely to end in two rounds than in three.

[5] See Dechenaux et al. (2012) for a recent comprehensive survey.

Malueg and Yates (2010) provide a simple and clean empirical study supporting the predictions of Klumpp and Polborn's (2006) model. Three laboratory experiments are modeled after Konrad and Kovenock's (2009) environment. In a best-of-seven series of all-pay auctions, Gelder and Kovenock (2014) also include a penalty for losing the series in some experimental treatments. As expected, they find that the presence of a large penalty induces the trailing contestant to exert more effort, resulting in narrower series winning margins than when losing penalties are smaller or absent. In best-of-three all-pay auctions, Mago and Sheremeta (2012) find that subjects significantly overbid in games two and three with the result that they go to game three more often than predicted and earn negative expected payoffs. Finally, Deck and Sheremeta (2012) report an experiment study in which players' observed effort exceeds the predicted levels.

The theoretical considerations present in the above-mentioned models do not apply to our setup because our game involves no costly effort and awards no prizes to the winners of the component games. Contests that require little effort from participants include penalty shootouts in soccer, free-throw competitions in basketball, spelling bees and game shows or, more generally, any contest in which the relevant skill or knowledge needed is acquired and rehearsed in advance and participation in the contest demands that the contestant merely apply this skill or knowledge in a timely manner. One advantage of contests without effort or intermediate prizes is that they enable us to test whether contestants are playing their precise equilibrium strategies. Indeed, in order to figure out the exact equilibrium of a two-outcome game without costly effort, intermediate prizes or other frictions, one needs to know neither players' risk preferences nor their cost functions. Since the Kingston-Anderson result assumes no effort on the part of the players and does not depend on risk preferences, it is a relatively straightforward task to design an environment appropriate for testing the theory. Furthermore, experimental results inconsistent with the theory would suggest more a failure of the theory than a misspecification of the experimental design. In this paper we report on an experiment that reliably tests the Kingston-Anderson predictions.

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 between rules in the quality of play at the game level: the frequency of equilibrium play is significantly higher in *winner* and *5-4* than in *alternating* and *loser*.

The paper is organized as follows. 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

To test the Kingston-Anderson equivalence of role-assignment rules, we require a game to be played in a best-of series format. We sought a two-player, zero-sum game with a unique pure-strategy Nash equilibrium that is neither too easy nor too difficult to solve. For this purpose, we choose the game "Duel".[6] The extensive-form version of Duel can be formalized as follows. 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 $N_1 = \{1,3,\ldots,19\}$ and $N_2 = \{2,4,\ldots,20\}$, respectively. At each node except 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, thus becoming 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 common knowledge, and increasing in n, meaning that by delaying 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

[6] Binmore (2007) provides a lively analysis of Duel. Ours is the first paper to bring this game to the laboratory.

fire first and thus end the game.

The "Bid-A-Note" phase of the popular, long-running TV game show "Name That Tune" is strategically equivalent to Duel. In Bid-A-Note, the game show host provides a clue to a song. Contestants then bid how many notes they require to name the song in question. The contestant who bids the fewer number of notes gets to hear this number of notes and then guesses the name of the song. If the guess is correct, the contestant wins the round; otherwise the opponent wins the round. The similarity between the play of Bid-A-Note and our game of Duel extends beyond the component game. Bid-A-Note was played in a best-of-five format in which the contestant who started the bidding alternated each round.

As defined above, Duel is a finite game of perfect information with chance moves. Moreover, since there are only two possible outcomes, namely, either Player 1 wins or Player 2 wins, and since the players' preferences over these outcomes disagree, Duel is a perfectly competitive game. The reason is that every pair of strategies induces a lottery over the two outcomes and the players' preferences over the set of such lotteries are diametrically opposed. Namely, Player 1 prefers one lottery over another if and only if the former assigns him a higher winning probability than the latter, and Player 2 prefers one lottery over another if and only if Player 1 prefers the latter to the former. Since all finite strictly competitive games of perfect information have a value, we conclude that Duel has a value as well. That is, there is a probability q^* and two strategies, s_1 for Player 1 and s_2 for Player 2, such that s_1 guarantees a winning probability for Player 1 of at least q^*, and s_2 guarantees a winning probability for Player 2 of at least $1-q^*$. Consequently, all Nash equilibria of Duel result in Player 1 winning with probability q^*. In fact, Player 1's equilibrium winning probability can be easily calculated. Let n^* be the first node n such that for some player $i \in \{1,2\}$,

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

Then, if n^* is one of Player 1's decision nodes $q^* = p_1(n^*)$, and if n^* is one of Player 2's decision nodes $q^* = 1 - p_2(n^*)$.

The reason for the above value of q^* is easy to understand when we assume that $p_i(n) + p_j(n+1) \neq 1$, for $i \neq j$ and for every node n. For in that case all Nash equilibria of Duel dictate that

Player $i = 1,2$ fires at the first node n_i such that

$$p_i(n_i) + p_j(n_i + 1) > 1, \qquad j \neq i, \tag{2}$$

but not earlier.[7] That is, in every Nash equilibrium the players plan to fire their shot at, and not earlier than, the first node at which the probability of hitting the opponent is higher than the probability of the opponent missing at the next node. To see this, denote by n^* and $n^* + 1$ the first nodes at which inequality (2) holds.[8] In particular, note that if players plan to fire at these nodes and not earlier, the gun will be shot at n^*. Consider now a pair of strategies such that at least one of the players does not behave as indicated. We argue that this cannot be an equilibrium. Indeed, if the gun is fired at $n < n^*$, then the shooter, say Player i, can profitably deviate by planning to shoot later at node $n+2$. He will then win with either probability $p_i(n+2)$ or $1 - p_j(n+1)$, both of which are greater than $p_i(n)$. If, on the other hand, the gun is fired at $n = n^*$, then the player who did not shoot, say Player j, must be the one deviating from equilibrium and planning to shoot later than at stage $n^* + 1$. Therefore, Player i can profitably deviate by delaying his shot and planning to shoot at $n+2$. He will then win with probability $p_i(n+2)$, which is greater than $p_i(n)$. Finally, if the gun is fired at $n > n^*$, the player who did not shoot, say Player j, can profitably deviate by shooting at $n-1$. He will win with probability $p_j(n-1)$ which exceeds $1 - p_i(n)$. As a result, the equilibrium outcome involves a gun being fired at the first node n such that inequality (2) holds.

In the experiment, we use the reduced normal-form representation of the game. Each contestant's action set consists of his ten decision nodes. That is, the action set of player 1 is $N_1 = \{1, 3, \ldots, 19\}$ and the action set of player 2 is $N_2 = \{2, 4, \ldots, 20\}$. 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 $n_1 \in N_1$ and $n_2 \in N_2$. Player 2's payoff function is $u_2(n_1, n_2) = 1 - u_1(n_1, n_2)$. Under the assumption that $p_i(n) + p_j(n+1) \neq 1$, for $i \neq j$ and for every node n, this game has a unique

[7] If $p_i(n) + p_j(n+1) = 1$ for some node n, both shooting and not shooting at that node is consistent with equilibrium and the probability that player 1 wins the game is the same in both cases.

[8] Since the functions p_1 and p_2 are increasing, once node n satisfies inequality (2) so do all subsequent nodes.

equilibrium, (n_1^*, n_2^*), corresponding to the first nodes identified above that satisfy (2). Clearly, the equilibrium action n_1^* guarantees that Player 1 wins with probability of at least q^*, and the equilibrium action n_2^* guarantees that Player 2 wins with probability of at least $1-q^*$, where q^* is the value of Duel identified above.

The game of Duel seems to be an apt choice for testing Kingston's theorem for a number of reasons. It is a constant-sum binary game whose reduced normal form has a unique equilibrium. Furthermore, this equilibrium is neither evident nor too difficult to figure out. The equilibrium consists of pure strategies, which has the benefit of not adding noise to the already random outcomes. We believe that Duel, having a unique pure-strategy equilibrium which is independent of the players' risk preferences, allows us to perform a reliable test of Kingston's result.

2.2 The series

A series consists of a sequence of multiple games of Duel. Two contestants, the *leader* and the *follower*, compete in $2k+1$ games of Duel with the series winner being the contestant who wins $k+1$ games. We consider a best-of-9 series, namely, a contest in which the first contestant to win five games wins the series.[9] 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 the specific role-assignment 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

[9] In the experiment (to be discussed in the next section), the winner of the series earns a fixed payment while the series loser receives nothing. No payments are awarded for victories in individual stage games.

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.[10] 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 $B(n, 5, q^*)$ where B stands for the binomial probability 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 $B(m, 4, 1 - q^*)$. Therefore, if the leader plays his equilibrium action in each game, he will win the series with a probability of at least

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

Similarly, if the follower adopts the equilibrium action in each of the component games, he will win the series with a probability of at least

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

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

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

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 mimics *winner* until the decisive duel. After the decisive duel, however, 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, this *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 that the leader wins the series. 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.

The Kingston-Anderson theorem provides us with a clear testable implication, namely that the proportion of series won by the leader is independent of the role-assignment rule. Additional implications can also be derived. 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 subsequent sections.

3 The experiment

3.1 Experimental design

To test the Kingston-Anderson 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 Section 2.2, will be referred to as *alternating*, *5-4*, *winner* and *loser*. In a between-subjects design, we conduct four sessions of each treatment. In each session, pairs of subjects play eight best-of-nine series of Duel preceded by one practice series.

In the instructions Duel is discussed in terms of a shootout. The shootout story was chosen to provide a concrete context to which subjects can relate. This was done to facilitate understanding.

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. The column "Series" indicates in which of the eight series the given parameter table ("Prm. Table") appeared. This ordering of parameter tables depends on the session number within the treatment and is the same for all treatments. 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. The equilibrium stages (not revealed to subjects) are highlighted in bold font.

A neutral framing would make the game more difficult to understand. See Cosmides (1989) for an early classic study that demonstrates how context substantially improves understanding in the Wason selection task.

The parameters for the nine series are displayed in Table 1. Each entry indicates 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 2.1) 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. For each set of game parameters, the unique equilibrium is highlighted in bold font (recall the equilibrium calculation derived in (2)).

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 parameterization 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 Player 1 does not deviate from his 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 a one-stage unilateral deviation by either player results 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). Presumably, the higher the cost of deviation, the easier it should be for subjects to arrive at their equilibrium actions. We will test this hypothesis in section 4.2.

At this point, a comment is in order about our choice of the game of Duel and the specific sets of game parameters. We believe we have chosen a game and game 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 the 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 game environment was not challenging enough.

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,[11] 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

[11] 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.

Within each game, subjects simultaneously choose at which stage to shoot. After both have entered their choices, the contestant who chose the earlier stage is the shooter and a random number generator determines whether he hits or misses his opponent. The pair of subjects then sees a screen summarizing the results of the game before proceeding to the next game.

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.[12] 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 new Israeli shekels (NIS) for each series they won plus a 30 NIS participation payment at the end of 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.[13] The entire experiment, including the instruction and payment phases, lasted up to two hours and 15 minutes.

[12] The instructions for the *alternating* treatment appear in the Appendix.

[13] At the time the experiments were conducted, $1 USD equaled approximately 3.5 NIS. The minimum wage in Israel is 22 NIS an hour.

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 2 displays the 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. Pairwise t-tests reveal that the average series lengths of any two treatments differ significantly from one another.

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*.

Despite these differences between treatments, the next four results demonstrate that the treatments are statistically indistinguishable from one another in the probability that a given contestant wins the series.

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

Result 1: The third-to-last row of Table 2 provides preliminary support for the hypothesis: the overall fraction of series won by the leader exhibits little variation, ranging from .545 (*winner*) to .580 (*alternating* and *5-4*). Regression analyses confirm the similarity across treatments. Table 3 displays the results from a linear probability model where the dependent measure is a binary variable for whether the leader won the series.[14] Regression (1) reports the estimates when all four set of parameters are pooled and parameter-table fixed effects and subject-specific random effects are included. Standard errors are clustered to account for any arbitrary correlation at the session level. The constant of .565 reflects the mean estimated percentage of series won by the leader in the omitted *alternating* treatment. The estimated coefficients on *5-4*, *winner* and *loser* are all close to and not significantly different from zero, implying that the fraction of series won by the leader in each of these treatments cannot be rejected as being the same as in *alternating*. Moreover, t-tests of coefficients show that none of the other pairs of treatments are significantly different from one another (p-values range from .50 to .88). To permit a time trend, a count variable for the series number $\{1,2,\ldots,8\}$ is included but not significantly different from zero in this or, unless otherwise indicated, any subsequent regression, nor does its inclusion affect the significance of the treatment indicators.[15]

In addition to testing the equivalence of the role-assignment rules on the pooled dataset, 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: Regressions (2), (3), (4) and (5) in Table 3 display separate regression results for each of the four parameter tables. None of the estimated coefficients on any of the treatment indicators

[14] 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 in this and the next subsection 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.

[15] All of our results are robust to alternative measures of a time trend, such as fixed effects for the series number or an indicator variable for the last four series of the experiment.

in any of the four regressions differs significantly from zero, again attesting to the similarity in the proportion of series won by the leader between *alternating* and each of the other three role-assignment rules. Moreover, pairwise t-tests of coefficients (12 in total) cannot reject the equality of the estimated coefficients on these three treatments in any of the four parameter tables. The lone exception is the weakly significant difference between the estimates on *5-4* and *winner* in parameter table 4 ($p = .09$).

Hypotheses 1 and 2 follow directly from Kingston's and Anderson's result, which states that the probability that the leader wins the series is independent of the role-assignment rule. We derive an analogous result regarding the winner of the first game. Concretely, it can be shown that 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.[16] 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 ($\sum_{n=0}^{4} \sum_{m=n-4}^{4} B(n,4,q^*)B(m,4,1-q^*)$) is independent of whether A is the leader or the follower.[17] Hypotheses 3 and 4 address this extension.

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

[16] 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.

[17] 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 follower and on the role-assignment rule.

Result 3: Initial evidence of the equivalence of the four role-assignment rules after game 1 of the series emerges from the second-to-last row in Table 2. The entries show that the proportion of series won by the winner of the series' first game varies narrowly from 64.6% (*loser*) to about 67.5% (*5-4* and *winner*) to 68.8% (*alternating*). Additional support for the assignment rules' post-game-one equivalence comes from the random-effects regression (6) in Table 4. The dependent variable equals 1 if the contestant – leader or follower – who won the series also won the first game of the series, and equals 0 if the series' winner did not win game 1 (or, equivalently, if the winner of game 1 lost the series). The estimated constant term of .663 in regression (6) reflects the estimated likelihood of winning the series if the contestant won the first game in the *alternating* treatment. None of the three treatment indicators is significantly different from zero (p-values range from .41 to .71), implying that in the *5-4*, *winner* and *loser* treatments the likelihood of winning the series conditional on winning game 1 does not differ significantly from that in *alternating*. Moreover, none of the pairwise t-tests of coefficients yield significant differences between the three former treatments (p-values from .56 to .92).

Analogous to regression (1), these estimated coefficients reflect averages over all four parameter tables. In regressions (7) – (10), we examine separately for each parameter table whether the contestant's likelihood of winning the series conditional on winning game 1 varies significantly across treatments. For parameter table 1, regression (7) shows that none of the three treatment indicators differs significantly from zero (p-values from .53 to 1); nor are any of the pairwise comparisons of the estimates significantly different from one another (p-values from .67 to .91). Regression (8) reveals that for parameter table 2 none of the estimated coefficients on the three treatment indicators differs significantly from zero (p-values from .74 to .85) and none of the pairwise t-tests reveals significant differences between these three treatments (p-values from .85 to 1). For parameter table 3, regression (9) again finds no significant differences between any of the four treatments, although the nearly 14 percentage-point difference between *5-4* and *winner* is close ($p = .14$). Finally, regression (10) shows that five of the six pairwise treatment comparisons are not significantly different from one another, while the 12.5 percentage-point gap between *5-4* and *winner* is weakly significant ($p = .07$). Note also that the estimated coefficient of .033 on the *series* count variable in (10) is weakly significant ($p = .07$), suggesting that the likelihood that the winner of game 1 goes onto win the series exhibits an upward trend over the course of the experiment. In summary, with 24 pairwise tests performed, we would expect 1.2 rejections of the null hypothesis at 5% or 2.4 rejections at 10%. Well in line with this expectation, we found a single rejection of the null at the 10% level. We conclude that the likelihood of the contestant who wins game 1 also wins the series is similar across treatments as hypothesized.

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-Anderson): 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: Hypothesis 3 compares the proportion of series won by the winner of game 1 across role-assignment rules without regard for whether the contestant played in the role of leader or follower. Hypothesis 4 asks whether leaders and followers differ in their proportion of series won conditional on winning game 1. Like Table 4, the dependent measure in the regression results that appear in Table 5 is a binary variable equal to 1 if the series winner also won game 1 and equal to 0 otherwise. The principal coefficient of interest is the estimate on “leader won game 1”, which equals 1 if the leader won the first game of the series and equals 0 if the follower won the series’ opening game. Thus, the interpretation of the estimated constant is the likelihood that the follower won the series given that he won the first game. The estimate on “leader won game 1” then addresses whether the leader’s probability of winning the series after winning game 1 differs from that of the follower. The estimate of -.004 ($p = .92$) in the overall regression (11) suggests that conditional on winning the first game of the series, the leader and follower are equally likely to go onto win the series. Regressions (12) – (15) display the results from each treatment separately.

None of the estimates on "leader won game 1" is significantly different from zero at conventional levels ($p = .18$ in *winner* and $p \geq .50$ in the other treatments). We conclude that the likelihood that the leader wins the series after winning the first game is not significantly different from that of the follower overall or in any of the four role-assignment rules.

Thus far, we have performed numerous regression analyses and statistical tests, all but two of which support the equivalence of the role-assignment rules in a given contestant's likelihood of winning the series – recall the weakly significant difference between *5-4* and *winner* in parameter table 4 in the proportion of series won by the leader (Result 2) and the weakly significant difference between *5-4* and *winner* in the proportion of series won by the winner of game 1 (Result 3). With 288 series played in each role-assignment rule and in each parameter table, 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 conduct the same regression analyses on series outcomes not predicted to be the same. As footnote 17 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 possible partial score at the beginning of game 3, Table 6 reports the estimates from a regression model identical to that in (1) of Table 3. The constant term reflects the estimated probability that the leader wins the series in the *alternating* treatment given the partial score after game 2. While all three constants in (16), (17) and (18) are significantly different from zero, they can be expected to and do diverge widely in their magnitudes from .146 when the leader trails zero games to two (0-2) to .824 when the leader is ahead two games to zero (2-0). Conditional on the 2-0 score, only the estimated coefficient of .165 on *loser* is significantly different from *alternating*. In addition, pairwise t-tests of coefficients reveal that the differences between the estimates on *5-4* and *loser* and between *winner* and *loser* are highly significantly different from one another.

The significant differences between *loser* and *5-4* and between *loser* and *winner* in the probability that the leader wins the series are both in accordance with the theory. To see this, note that a 2-0 score in *5-4* and *winner* implies that the leader has played both games as Player 1, leaving him with at most three games in the advantaged role, whereas the follower still has all four games to play as the advantaged player. By contrast, a 2-0 score in *loser* means that the leader and the follower have each used up one play in the advantaged role of Player 1, leaving the leader with up to four and the follower with up to only three games in the advantaged role. Consequently, ahead 2-0, the leader is more likely to win the series in *loser* than in *5-4* or *winner* – a prediction borne out by the significantly higher estimated coefficient of .165 on the *loser* treatment indicator. At the same time, the significant difference between *alternating* and *loser* was not anticipated by the theory. In both of these treatments, a 2-0 partial score implies that the leader and the follower have each used up the same number of games – one – in the advantaged. Thus, the probability that the leader goes onto win the series after leading 2-0 ought to be the same in these two treatments.

When the partial score is 1-1, the theory predicts that, having played only one game as Player 1, the leader in *alternating* is more likely to win the series than the leader in *5-4* who already used up two games as Player 1. The highly significant estimated coefficient of -.145 on the *5-4* treatment indicator in (13) suggests that, in accordance with the theoretical prediction, the leader in *alternating* is 14.5 percentage points more likely to win the series than his *5-4* counterpart. Regression (17) also reveals that the leader is significantly more likely to win the series in *alternating* than in *winner* or in *loser*. The theory predicts this result in series in which the first game was won by the leader in *winner* and the follower in *loser*; otherwise, no difference between these pairs of treatments is predicted.

Finally, when the leader trails the series zero games to two (0-2) – a rarer occurrence than the other two partial scores – he is expected to be less likely to win the series in *5-4* and *loser* (where he has already played twice as Player 1) than in *alternating* and *winner* (where the follower also played once as Player 1). None of these predictions is borne out by the data. In fact, owing in part to the paucity of observations associated with this unlikely partial score, there are no significant differences between any treatment pairs.

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 often 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 the Kingston-Anderson equivalence result since the underlying probabilities of winning a game and the overall series would be unobservable. No further test of the equivalence 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 the Kingston-Anderson theorem.

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

Result 6: The third-to-last row of Table 2 displays the realized fractions of series won by the leader and the corresponding theoretical probability averaged over all four game parameters for each of the treatments. Comparing treatments, the realized fractions of series won by the leader span only 3.5 percentage points from .545 to .580 and all fall within two percentage points of the expected fraction of .562. Wald tests of the estimated coefficients in regression (1) in Table 3 reveal that none are significantly different from .562 (p-values from .50 to .96). Turning to the estimates in the separate regressions (2) – (5) for each parameter table, none of the estimates for parameter table 1 differ significantly from the theoretical value of .568 (p-values from .78 to .91). Nor can we reject at conventional significance levels the equivalence of the estimated treatment effects and the theoretical expectation (displayed in the bottom row of Table 3) for any of the remaining three parameter tables (the lowest p-value is .16 for *winner* in parameter table 4). In short, all 20/20 Wald tests fail to reject Hypothesis 6.

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, Binomial tests 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 their equality 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.

Result 7: The second-to-last row of Table 2 displays the realized fractions of series won by the leader averaged over all four game parameters for each of the treatments. These fractions range from .646 *loser* to .688 *alternating* compared to the theoretical prediction of .653. Wald tests to determine whether the estimated constant in regression (6) of Table 6 as well as the constant plus, in turn, each of the estimated coefficients on the treatment indicators fail to reject the null hypothesis that each of these fractions equals .653 (p-values from .61 to 1). Further support for Hypothesis 7 comes from comparable Wald tests for each of the four parameter tables. None of the 16 estimated expressions from regressions (7) – (10) differs significantly from the corresponding theoretical prediction (shown in the last row of Table 4).

Until now, our focus has been on comparing treatments to one another and to the theoretical point predictions on the basis of series-level outcomes. 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 (i.e., observed choices) at the game level where the theory's predictive power reveals its first cracks.

4.2 Game-Level Results

The last column of the first row in Table 7 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 7 highlights that, according to several distinct measures, play is closer to equilibrium 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%).[18] 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

[18] 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.

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 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.[19]

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

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

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. Parameter-table fixed effects are captured by γ_r. Finally, ε represents the idiosyncratic error term. Standard errors are clustered by subject, taking into account the correlation in the error terms over the games and series within a subject. Table 8 presents the results.

Regression (19) displays the marginal effects from three of the four treatments. The constant of .461 reflects the mean percentage of equilibrium play in the omitted *alternating* treatment; in *loser* this fraction is not significantly different from that in *alternating* ($p = .46$), 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 ($p = .77$). 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

[19] 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.

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 (20) 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 (21) indicates that moving from a low-cost to a high-cost parameter table raises the frequency of equilibrium play 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.

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 draw the following analogy with 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 (21) 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 corresponding 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 (22).

Within a series, the likelihood of playing the equilibrium action increases by 1.1 percentage points from one game to the next ($p < .01$). From series to series, the proportion of equilibrium play rises by 2.3 percentage points ($p < .01$). On the other hand, the point estimate of .027 for the importance-of-the-game variable is not significantly different from 0 ($p = .31$).[20] 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.

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 (23) 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 (23), the highly significant coefficients of .117 and .132 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

[20] 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 significantly different from zero in each treatment. The importance-of-the-game variable is not significant in three of the treatments (p-values are .40, .90 and .93) and only marginally significant ($p = .06$) in *winner*.

importantly, the coefficients on the treatment-follower interaction terms are all tiny and not significantly different from zero.[21] 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.

Regressions parallel to those in Table 8 appear in Table 9 with the dependent variable y_{igr} in equation (5) 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, *5-4* and *alternating* are the only two treatments even weakly significantly different from one another ($p = .09$ in each of (24) – (27) and $p = .04$ in (28) in Table 9). No other treatments can be rejected as being similar at conventional levels of significance. Regression (25) shows that Player 1 is significantly more 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 (26). Similar to regression (22), (27) 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 (28) 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 nearly two percentage points less often than the leader. These results in (28) are robust to the inclusion of the controls from previous regressions.

[21] This identical result (not included in the table) is obtained if we include some or all of the controls from the 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 methods of role assignment 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.

In a recent paper, Mago et al. (2013) report the results of an experiment on a symmetric best-of-three Tullock contest, i.e., where the probability of winning each game depends on the relative levels of effort invested by the contestants. They distinguish between strategic and psychological momentum. The game exhibits strategic momentum when the equilibrium probability that the winner of the first game also wins the second game is higher than 1/2. Play exhibits psychological momentum if the probability that the winner of the previous game also wins the current game is higher than the corresponding equilibrium probability. Their results show that play is consistent with strategic momentum, while little evidence is found for psychological momentum. In our design of Duel, dependence of the frequency of the equilibrium choice on partial scores or on the importance of the game would be evidence of psychological momentum. Similar to Mago et al. (2013), however, the results reported in Section 4.2 suggest the absence of psychological momentum in our implementation of Duel.

By definition, psychological momentum can arise only as a result of suboptimal play by at least one of the players, which in turn 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 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] Cosmides, Leda. “The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task,” *Cognition* 31:3, 187–276 (1989).

[7] Dechenaux, Emmanuel, Kovenock Dan and Sheremeta, Roman. “A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments,” mimeo (2012).

[8] Deck, Cary and Roman M. Sheremeta. “Fight or flight? Defending against sequential attacks in the game of Siege,” *Journal of Conflict Resolution* 56:6, 1069–1088 (2012).

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

[10] Fréchette, Guillaume R. “Session-Effects in the Laboratory,” *Experimental Economics* 15:3, 485–498 (2012).

[11] Gelder, Alan and Dan Kovenock. “Fight or Surrender: Experimental Analysis of Last Stand Behavior,” mimeo (2014).

[12] Harris, Christopher and John Vickers. “Racing with Uncertainty,” *The Review of Economic Studies* 54:1, 1–21 (1987).

[13] Irfanoglu, Zeynep B., Shakun D. Mago and Roman M. Sheremeta. “Sequential versus Simultaneous Election Contests: An Experimental Study,” Mimeo (2010).

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

[15] Klumpp, Tilman and Mattias K. Polborn. “Primaries and the New Hampshire Effect,” *Journal of Public Economics* 90, 1073-1114 (2006).

[16] Kocher, Martin G., Marc V. Lenz and Matthias Sutter. “Psychological Pressure in Competitive Environments: New Evidence from Randomized Natural Experiments,” *Management Science* 58:8, 1585–1591 (2012).

[17] Konrad, Kai A. and Dan Kovenock. "Multi-battle Contests," *Games and Economic Behavior* 66, 256–274 (2009).

[18] 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).

[19] Mago, Shakun D. and Roman M. Sheremeta. "Multi-Battle Contests: An Experimental Study," Mimeo (2012).

[20] Mago, Shakun D., Roman M. Sheremeta and Andrew Yates. "Best-of-Three Contest Experiments: Strategic versus Psychological Momentum," *International Journal of Industrial Organization* 31(3), 287–296 (2013).

[21] Malueg, David A., and Andrew J. Yates. "Testing Contest Theory: Evidence from Best-of-Three Tennis Matches," *Review of Economics and Statistics* 92(3), 689–692 (2010).

[22] 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).

[23] 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).

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

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

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

[27] Zizzo, Daniel J. "Racing with Uncertainty: A Patent Race Experiment," *International Journal of Industrial Organization* 20, 877-902 (2002).

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 2 – Summary Statistics

	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
Fraction of series won by Leader	.580	.562	.580	.562	.545	.562	.552	.562
Fraction of series won by winner of game 1	.688	.653	.674	.653	.677	.653	.646	.653
Total Obs.	288	288	288	288	288	288	288	288

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 followed by the fraction of series won by the leader and the fraction of series won by the contestant who won game 1 of the series.

Table 3 – Frequency that Leader won the series

Regression	(1)	(2)	(3)	(4)	(5)
Parameter Table	All	1	2	3	4
constant	.565 (.053)	.580 (.087)	.582 (.073)	.681 (.084)	.596 (.122)
5-4	0 (.053)	0 (.102)	-.014 (.116)	.014 (.068)	0 (.114)
winner	-.035 (.035)	-.028 (.087)	.069 (.061)	-.056 (.080)	-.125 (.117)
loser	-.028 (.050)	-.042 (.057)	-.014 (.064)	.028 (.164)	-.083 (.127)
series	-.006 (.007)	-.009 (.018)	.006 (.012)	-.022 (.020)	-.003 (.012)
R^2	.008	.003	.006	.012	.012
Obs	1152	288	288	288	288
Parameter table fixed effects	YES	NO	NO	NO	NO
Theoretical Probability	.562	.568	.543	.581	.553

Notes: 1. Linear probability model to test hypotheses 1, 2 and 6. Subject random effects are included in regression (1) to model possible subject heterogeneity.
2. Dependent variable: binary variable for whether the leader won the series in all parameter tables combined (regression (1)) and by parameter table, (2) - (5).
3. Independent variables: indicators for the *5-4*, *winner* and *loser* treatments and the series number according to its order of appearance in the session.
4. Standard errors are clustered to account for any arbitrary correlation at the session level.
5. The last row "Theoretical Probability" indicates the expected probability that the leader wins the series, assuming both contestants play their equilibrium strategies.
6. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 4 – Frequency that Contestant who won the first game also won the series

Regression	(6)	(7)	(8)	(9)	(10)
Parameter Table	All	1	2	3	4
constant	.663 (.047)	.672 (.081)	.693 (.074)	.645 (.073)	.571 (.111)
5-4	-.014 (.028)	0 (.088)	-.028 (.080)	-.097 (.084)	.069 (.065)
winner	-.010 (.028)	-.014 (.101)	-.014 (.073)	.042 (.066)	-.056 (.084)
loser	-.042 (.051)	-.042 (.065)	-.028 (.088)	-.028 (.136)	-.069 (.106)
series	.008 (.006)	.005 (.016)	-.006 (.011)	.005 (.013)	.033* (.017)
R^2	.005	.002	.002	.012	.038
Obs	1152	288	288	288	288
Parameter table fixed effects	YES	NO	NO	NO	NO
Theoretical Probability	.653	.655	.644	.662	.648

Notes: 1. Linear probability model to test hypotheses 3 and 7. Subject random effects are included in regression (6) to model possible subject heterogeneity.
2. Dependent variable: binary variable for whether the contestant who won the first game of the series also won the series, in all parameter tables combined (regression (6)) and by parameter table, (7) - (10).
3. Independent variables: indicators for the *5-4*, *winner* and *loser* treatments and the series number according to its order of appearance in the session.
4. Standard errors are clustered to account for any arbitrary correlation at the session level.
5. The last row "Theoretical Probability" indicates the expected probability that the leader wins the series after having won game 1 of the series, assuming both contestants play their equilibrium strategies.
6. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 5 – Frequency that Contestant who won the first game also won the series

Regression	(11)	(12)	(13)	(14)	(15)
Treatments	All	Alternating	5-4	Winner	Loser
constant	.649 (.054)	.633 (.074)	.697 (.082)	.644 (.157)	.639 (.109)
leader won game 1	-.004 (.037)	-.024 (.035)	.058 (.101)	-.083 (.062)	.027 (.084)
series	.008 (.006)	.017 (.014)	-.011 (.007)	.022*** (.006)	-.000 (.009)
R^2	.004	.010	.036	.021	.001
Obs	1152	288	288	288	288
Parameter table fixed effects	YES	YES	YES	YES	YES

Notes: 1. Linear probability model to test hypothesis 4.
2. Dependent variable: binary variable for whether the contestant who won the first game of the series also won the series, in all treatments combined (regression (11)) and separately for each treatment, (7) - (10).
3. Independent variables: *leader won game 1*: indicator variable equal to 1 if the leader won game 1 of the series and equal to 0 if the follower won game 1 and the series number according to its order of appearance in the session.
4. Standard errors are clustered to account for any arbitrary correlation at the session level.
5. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 6 – Frequency that Leader won series conditional on partial score after game 2

Regression	(16)	(17)	(18)
Partial Score	2-0	1-1	0-2
constant	.824	.555	.146
	(.092)	(.080)	(.063)
5-4	-.029	-.145***	.103
	(.078)	(.044)	(.081)
winner	.042	-.201***	.024
	(.091)	(.047)	(.067)
loser	.165**	-.117*	-.013
	(.075)	(.065)	(.075)
series	-.015	-.003	-.013
	(.011)	(.010)	(.011)
R^2	.037	.026	.114
Obs	409	575	168
Parameter table fixed effects	YES	YES	YES

Notes: 1. Linear probability model to test hypothesis 5.
2. Dependent variable: binary variable for whether leader won the series, with a separate regression for each partial score after game 2.
3. Independent variables: same as Table 3: indicators for the *5-4*, *winner* and *loser* treatments and the series number according to its order of appearance in the session.
4. Standard errors are clustered to account for any arbitrary correlation at the session level.
5. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 7 – Measures of equilibrium play

Equilibrium Play Measure	Alternating	5-4	Winner	Loser	Overall
	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)
% eq'm action	0.494 (0.500)	0.613 (0.487)	0.626 (0.484)	0.522 (0.500)	0.561 (0.496)
% eq'm action + 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 pair of actions	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 pair of actions refers to the percentage of games in which both paired contestants chose their equilibrium actions.

Table 8 – Regressions on the frequency of equilibrium play

Regression	(19)		(20)		(21)		(22)		(23)	
5-4	0.119	***	0.119	***	0.116	***	0.118	***	0.117	***
	(0.039)		(0.039)		(0.039)		(0.039)		(0.042)	
winner	0.131	***	0.131	***	0.125	***	0.134	***	0.132	***
	(0.040)		(0.040)		(0.040)		(0.040)		(0.043)	
loser	0.029		0.029		0.032		0.029		0.031	
	(0.038)		(0.038)		(0.038)		(0.038)		(0.041)	
leader			-0.007							
			(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	***	0.027			
					(0.025)		(0.026)			
series							0.023	***		
							(0.002)			
game							0.011	***		
							(0.002)			
alternating*follower									0.007	
									(0.020)	
5-4*follower									0.011	
									(0.015)	
winner*follower									0.005	
									(0.017)	
loser*follower									0.003	
									(0.017)	
constant	0.461		0.463		0.444		0.300		0.457	
	(0.027)		(0.027)		(0.027)		(0.028)		(0.029)	
R^2	.016		.016		.016		.030		.016	
Obs	17092		17092		17092		17092		17092	
Parameter table fixed effects	YES		YES		NO		YES		YES	

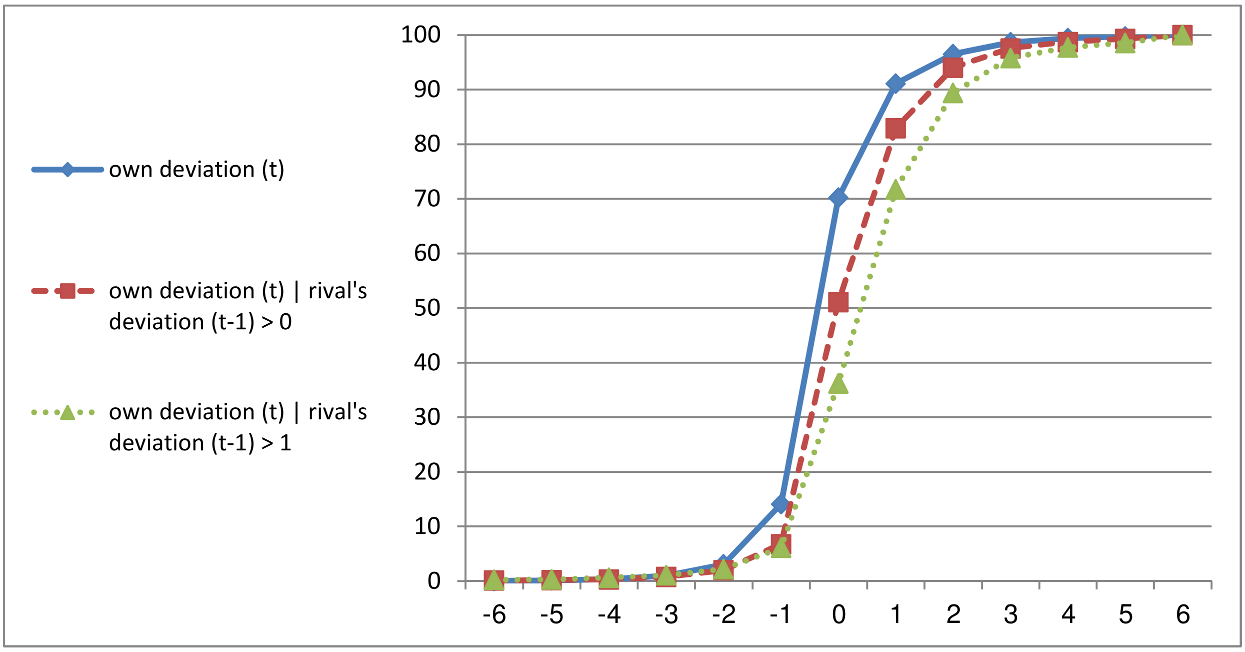
Notes: 1. Linear probability model to test hypothesis 8.
2. Dependent variable: binary variable for whether subject *i* in game *g* of series *r* played the equilibrium action.
3. Independent variables: indicators for *5-4*, *winner* and *loser* treatments; indicators for whether subject *i* was the 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; fixed effects for parameter tables are included where noted.
4. Standard errors are clustered by subject to account for any arbitrary correlation in the error terms across games and series within a subject.
5. * 10 percent significance level, ** 5 percent significance level, *** 1 percent significance level.

Table 9 – Regressions on the frequency of equilibrium play or within one stage of it

Regression	(24)		(25)		(26)		(27)		(28)	
5-4	0.045	*	0.045	*	0.044	*	0.045	*	0.056	**
	(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		0.017	
	(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					0.020		0.000			
					(0.016)		(0.016)			
series							0.011	***		
							(0.002)			
game							0.003	**		
							(0.001)			
alternating*follower									0.005	
									(0.016)	
5-4*follower									-0.017	*
									(0.009)	
winner*follower									-0.009	
									(0.011)	
loser*follower									0.013	
									(0.014)	
constant	0.864		0.840		0.854		0.798		0.862	
	(0.019)		(0.020)		(0.020)		(0.016)		(0.020)	
R^2	0.003		0.009		0.009		0.010		0.004	
Obs	17092		17092		17092		17092		17092	
Parameter table fixed effects	YES		YES		NO		YES		YES	

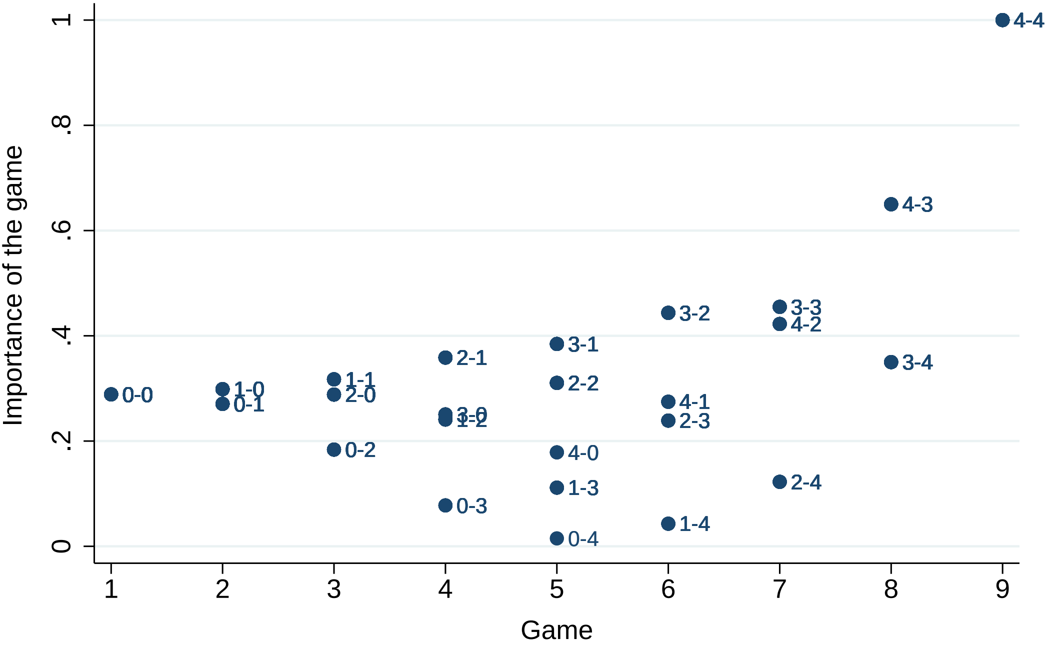
Notes: 1. Dependent variable: binary variable for whether subject *i* in game *g* of series *r* played the equilibrium action or within one stage of the equilibrium.
2. See Table 8 for further explanations.

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