Minimax Play in Armageddon Chess

Working Paper #0121

Miguel A. Fonseca, Todd Kaplan, Andrew Wooders, and John Wooders

NYU Abu Dhabi
November 2026

جامعة نيويورك أبوظبي
NYU ABU DHABI

Minimax Play in Armageddon Chess*

Miguel A. Fonseca† Todd Kaplan‡ Andrew Wooders§

John Wooders¶

November 1, 2026

Abstract

We use data on bids and game outcomes from Armageddon chess in high-stakes professional chess tournaments to undertake the first test using field data of the Minimax Hypothesis, and its generalization to Nash equilibrium, when the solution is in pure strategies. Under Armageddon rules, Black wins either by winning or drawing, while White must win outright. In the game we study, players bid time for the right to play Black. The player who makes the smaller bid plays Black, with the starting time on their clock equal to their bid; their opponent plays White with a full clock. We find that bids and

*We are grateful to Erik Allebest for providing us with Chess.com data on Bidding Armageddon, to Michael Brancato for advice and useful discussions, and to Jonathan Rosenthal for providing us access to a research-ready version of his chess engine, Winter. We are grateful to Gabriel Moller, Adilbek Tursynbek, and Scott Vincent for expert research assistance. We are grateful to Navin Kartik for comments.

†Department of Economics, University of Exeter (m.a.fonseca@exeter.ac.uk).

‡Department of Economics, University of Haifa and Department of Economics, University of Exeter (t.r.kaplan@exeter.ac.uk).

§Kellogg Marketing Department, Northwestern University (andrew.wooders@kellogg.northwestern.edu).

¶Center for Behavioral Institutional Design and Division of Social Science, New York University Abu Dhabi (john.wooders@nyu.edu). Wooders gratefully acknowledges financial support from Tamkeen under the NYU Abu Dhabi Research Institute Award CG005.

game outcomes are consistent with minimax and hence Nash play: each player is equally likely to play Black or White, and a player's probability of winning the game is the same whether playing Black or White. In this latter respect, Armageddon bidding fairly divides the probability of winning a match. We show that players sacrifice time, and hence playing accuracy, in exchange for the advantage of playing Black. In the data, these two factors exactly balance, as predicted by theory.

1 Introduction

Empirical tests of game theory require sufficient knowledge of the game – its actions and payoffs – for the theory to generate testable hypotheses. This condition is rarely satisfied by data generated in the field. A notable exception is professional tennis, where the contest between server and receiver for a point is a two-player constant-sum game. While theory does not identify the Nash equilibrium mixture, it does yield several testable hypothesis, one being that the server's probability of winning the point is the same for every action in the support of their equilibrium mixture.

The present paper provides the first field test of the Minimax Hypothesis, and of its generalization to Nash equilibrium when the minimax solution is in pure strategies. It exploits data from Bidding Armageddon games used to break ties in matches in professional chess tournaments. In Bidding Armageddon, two players simultaneously bid amounts of time, and they then play a timed game of chess in which Black has "draw odds," i.e., Black is declared the winner if the game ends in a draw. The player who makes the smaller bid chooses whether to play Black or White, with an amount of time on their clock equal to their bid, while the other player has a full clock. If the bids are the same, the chooser is selected at random. The winner of the Armageddon game wins the match.[1]

When formulating bids, players trade off the benefit of playing Black with draw odds against the cost of having less time on their clocks. The equilibrium bid is determined by an indifference condition. Specifically, in equilibrium, both players are simultaneously indifferent between playing Black with an amount of time on their clock equal to the equilibrium bid and playing White with a full clock.

Theory does not predict the bids or how they depend on match-specific

[1] In a famous match in 2023 between Magnus Carlsen and Hikaru Nakamura, ranked first and second in the world, Nakamura bid 8 minutes and 59 seconds and Carlsen bid 8 minutes and 58 seconds. Carlsen won by a single second and took Black; Nakamura played White with 15 minutes on his clock. Carlsen won with a draw.

characteristics, such as the players' abilities. Nevertheless, theory makes five sharp, testable predictions. First, the "stronger" player, i.e., the player with the higher Elo rating, plays White with probability .5. Second, player ratings of Black are drawn from the same distribution as player ratings of White. Third, and most important, color does not predict match outcomes: a player wins with the same probability whether playing Black or White. Fourth, White wins with probability 1/2. Last, both players make the same bid.

We test these predictions using data from all the Bidding Armageddon games hosted by Chess.com between August 2022 and May 2025, involving 523 games and 203 players across 24 high-stakes tournaments. All the players in our dataset are titled, with 91% being Grandmasters, with an average rating of 2669. These players represent the elite of international competitive chess. For each Bidding Armageddon game, we observe the players' ratings, their bids, the color assignment, and the Armageddon game outcome. We have, in addition, for every non-Armageddon game in the same tournament bracket in which an Armageddon game took place, the identities of the players, their color assignments, the sequence of moves, and the game outcome. Of particular interest will be the those games played in a match that led up to a Bidding Armageddon game.

We find that, although bids are not equal, they are close. The data supports all the other theoretical predictions. Importantly, the data supports the equilibrium condition that a player's winning probability is independent of color assignment. This feature is the primary motive for the development and introduction of Armageddon bidding.

We use data on time spent per move and move quality, as measured by a chess engine, to look inside the black box that determines game outcomes.[2] We show that the different incentives of Black and White for the game to end in a draw cause Black to make moves that increase the likelihood of a draw, while White does the opposite. Moreover, except at the beginning of a game,

[2] A chess engine is software that evaluates board positions and searches possible move sequences to choose a strong move.

Black uses less time per move, resulting in lower-quality play than White. We find that the cost of playing Black with less time exactly offsets the advantage of having draw odds, leaving players indifferent between playing Black and White, as predicted by theory.

Data from Bidding Armageddon is ideal for testing game theory for several reasons. First, the rules of the game are simple and unambiguous. Second, the game has exactly two outcomes – one player or the other wins – and is strictly competitive. Hence, neither risk attitudes nor social preferences enter payoffs. No auxiliary assumptions, such as risk neutrality, are required, yielding a clean test of the theory. Third, the stakes in Bidding Armageddon are substantial, with prize pools in our data ranging from $5,000 to $500,000, and individual prizes up to $200,000, ensuring that players have strong incentives to play optimally. Fourth, the players are highly experienced and know each other's play. They know the trade-off between the benefit of playing Black with draw odds – measured in terms of the probability of winning – and the cost of having less time on their clock. In other words, the players know the payoff function. Finally, even though all the players in our dataset are elite players, there are nonetheless significant differences in their ratings and playing styles, thus providing a test of the Minimax Hypothesis, and equivalently Nash equilibrium, in a rich environment.

Equilibrium is often motivated as the outcome of players learning through repeated play. In professional tennis, for example, a player makes about 150 first serves over the course of a single five-set match. By contrast, our dataset typically contains only a handful of Bidding Armageddon games for each player. Thus, it is implausible that equilibrium behavior in Bidding Armageddon emerges through reinforcement or belief-based learning, since this tiebreak format only gained prominence in 2022 and such tiebreaks are relatively uncommon.

The theoretical predictions we test do not depend on players selecting mutual best responses à la Nash. It seems more plausible, instead, that equilibrium emerges from maximin reasoning: each player chooses a bid that max-

imizes their minimum payoff against any bid their opponent might choose. The finding that behavior and outcomes conform closely to theoretical predictions therefore provides support for the Minimax Hypothesis.

Fair Division

Bidding Armageddon is a real-world solution to a fair division problem: How to fairly divide, between two players, a unit probability of winning a chess match. As recently as 1983, a Candidates match, part of the qualification process for the World Chess Championship, was decided by a spin of a roulette wheel. When first introduced in the late 1990s, Armageddon games had fixed time controls for each color. In "blitz" games, White had 5 minutes on their clock, while Black had 4 minutes. In the "rapid" format, White had 10 minutes, while Black had 7 or 8 minutes. The color assignment was determined by drawing lots.[3] An undesirable feature of the random assignment of color is that luck plays a consequential role if one color is advantaged over another. In Bidding Armageddon, by contrast, at the equilibrium bid, the players are indifferent between colors. The probability that a player wins depends only on skill. Luck plays no role.

Bidding Armageddon can be viewed as a generalization of the classic Nash (1953) demand game. Suppose that a dollar is to be divided. In Nash's game, each player simultaneously demands a share. If the shares sum to less than one, then each player receives their demand. Consider the variation in which the player who demands the smaller share receives their demand and the other player receives the residual share.[4] This game has a unique Nash equilibrium (which coincides with the Minimax solution). In equilibrium each player demands exactly a one-half share. We will show that Bidding Armageddon, in which the player who bids/demands the smaller amount of time gets to choose Black or White, is a generalized version of this variation.

[3] An arbiter hides a white pawn in one hand and a black pawn in the other. One player chooses a hand; the color revealed is the color they play.

[4] In the event that both players make the same demand, a player chosen at random receives their demand, and the other player receives the residual.

Asymmetric contests, in which one side is favored over the other, are common even when the players themselves are symmetric. In chess, Black is advantaged when it has draw odds. In a Texas Shoot-out, a mechanism commonly used to dissolve business partnerships, it is advantageous to be the second mover (the "chooser") when participants have independent private values. Our results illustrate that allowing players to bargain over roles can successfully resolve such asymmetries.

Related Literature

Von Neumann's (1928) Minimax Theorem for two-player zero-sum games and Nash's (1950) generalization to the notion of Nash equilibrium for n player games with general payoffs, are the foundation of modern game theory and our understanding of strategic behavior. Mixed-strategy Nash equilibrium has not performed well in the many laboratory experiments using human subjects carried out to test the theory (see O'Neill (1987, 1991) and Brown and Rosenthal (1990)). In contrast, following Walker and Wooders (2001), there is now substantial evidence that the behavior of elite sports professionals, in tennis and soccer, is largely consistent with theory.[5] For results on tennis, see Hsu, Huang, and Tang (2007), Paserman (2023), Gonzalez-Diaz, Gossner, and Rogers (2012), Gauriot, Page, and Wooders (2023), and Anderson, Rosen, Rust, and Wong (2025). For results on soccer, see Chiappori, Levitt, and Groseclose (2002), Palacios-Huerta (2003), Levitt, List, and Reiley (2010), and Wooders (2010). All of these papers test the Minimax Hypothesis when the solution is in mixed strategies, while the present paper tests the Hypothesis when the solution is in pure strategies.

Our paper also relates to studies of strategic reasoning by chess players. Using game data from lichess.org, a popular online chess platform, Salant and Spenkuch (2026) study chess endgames in which six or fewer pieces remain on the board. The backwards induction solutions for such games

[5] The evidence from sports is inconsistent with theory in one respect. Walker and Wooders (2001) show that there is serial correlation in the direction of serve in professional tennis. Gauriot, Page, and Wooders (2023) show that behavior conforms more closely to theory for men than women, and more closely for top women than non-top women.

are known. They report that when players have a meaningful choice (i.e., the game has at least one winning move and at least one drawing or losing move), then a player wins in only 74% of such games. In other words, a non-trivial fraction of chess players do not follow a backwards induction solution, even when playing chess. Their central finding is that the evidence favors satisficing rather their model of noisy maximization. Levitt, List, and Sadoff (2011) study backward induction by tournament chess players in two Race-to-100 games and the Centipede game. Even in the Race-to-100 games, which provide a clean test of backward induction, only 15 of 204 players (7.3%) follow the backward induction solution. In the Centipede game, only 3.9% of chess players stop at the first decision node.

There is a long history of laboratory experiments on pure-strategy Nash equilibrium play in normal form games. Indeed, Flood and Dresher conducted what may be the first experiment on strategic decision-making, in 1950, in an experiment that introduced the now-celebrated Prisoner's Dilemma game (see Roth (1993)). Nash equilibrium performs poorly in one-shot normal form games. Costa-Gomes and Weizsäcker (2008), for example, study 14 two-player 3×3 normal-form games with unique pure-strategy Nash equilibria. The choice frequency for the Nash actions is 36.1%, which is only slightly above random play. By contrast, Nash equilibrium often predicts well when, with repetition, players have the opportunity to learn. Nagel (1995) shows that play converges to the Nash equilibrium in the Beauty Contest game.

The consistent failure of pure-strategy Nash equilibrium in laboratory experiments on one-shot games, where there is no scope for learning, has led to the creation of new models, such as level-k reasoning, designed specifically to explain initial play (see Stahl and Wilson (1994, 1995) and Nagel (1995)). This paper provides the first field test of the Minimax Hypothesis and Nash equilibrium, when the minimax solution is in pure strategies. Our results show that bids and game outcomes conform remarkably closely to theory.

2 Theory: Bidding Armageddon

In the model, we assume that the player with the lower bid chooses Black, and hence a player's only decision is their bid.[6] Let $\pi(t, r_1, r_2)$ denote the probability that Black wins when (i) Black has t minutes on their clock and White has T minutes ($t \leq T$), (ii) r_1 is the rating of the player assigned to Black, and (iii) r_2 is the rating of the player assigned to White. If player i's probability of winning as Black is $\pi(t, r_i, r_j)$, then player i's probability of winning as White is $1- \pi(t, r_j, r_i)$. Since the game has only two outcomes, it is without loss of generality to take each player's payoff to be the probability that the player wins.

We make three mild assumptions about winning probabilities. First, Black loses for certain when there is no time on Black's clock, i.e.,

$$\pi(0, r_1, r_2) = 0 \text{ for all } (r_1, r_2).$$

Second, $\pi(t, r_1, r_2)$ is continuous and strictly increasing in t. In other words, Black's winning probability is increasing in the time on Black's clock, when White has a full clock T. Finally, we assume that it is advantageous to play Black when Black and White both have full clocks, i.e.,

$$\pi(T, r_1, r_2) > 1 - \pi(T, r_2, r_1) \text{ for all } (r_1, r_2).$$

In the regular (non-Armageddon) games in our data set, where both players have a full clock, Black wins or there is a draw in, respectively, 27.5% and 36.3% of the games. Hence having draw odds conveys a significant advantage to Black.

To identify the minimax solution, it is useful to consider Figure 1(a) below which shows Player 1's payoff when playing Black and when playing White,

[6] In the data, the winning bidder chooses Black almost without exception. In the game in which the winning bidder chooses their color, the first two assumptions of our model guarantee the existence of a unique Nash equilibrium, while the third ensures that in equilibrium the winning bidder chooses Black.

when Black has t minutes on their clock. Consider the bid t_1' shown on the horizontal axis. If Player 1 bids t_1', then Player 1's worst-case payoff is when Player 2's bid is greater than t_1', in which case Player 1 is assigned Black and obtains a payoff of $\pi(t_1', r_1, r_2)$, shown as the red dot directly above t_1'. This payoff is less than any payoff that Player 1 might obtain as White. In particular, if Player 2 bids less than t_1', then Player 1 is assigned White and has a payoff of at least $1 - \pi(t_1', r_2, r_1)$, shown as the green dot directly above t_1', which exceeds $\pi(t_1', r_1, r_2)$.

If Player 1 bids t_1'', then the worst-case payoff for Player 1 is when Player 2 bids just below t_1'', in which case Player 1 is assigned White and has a payoff of approximately $1 - \pi(t_1'', r_2, r_1)$, shown as the red dot above t_1''. This payoff is less than the payoff Player 1 would obtain as Black. In particular, if Player 2 bids more than t_1'', then Player 1 is assigned Black and has a payoff of $\pi(t_1'', r_1, r_2)$.

Figure 1(b) shows, in red, Player 1's worst-case (i.e., minimum) payoff for each bid t.

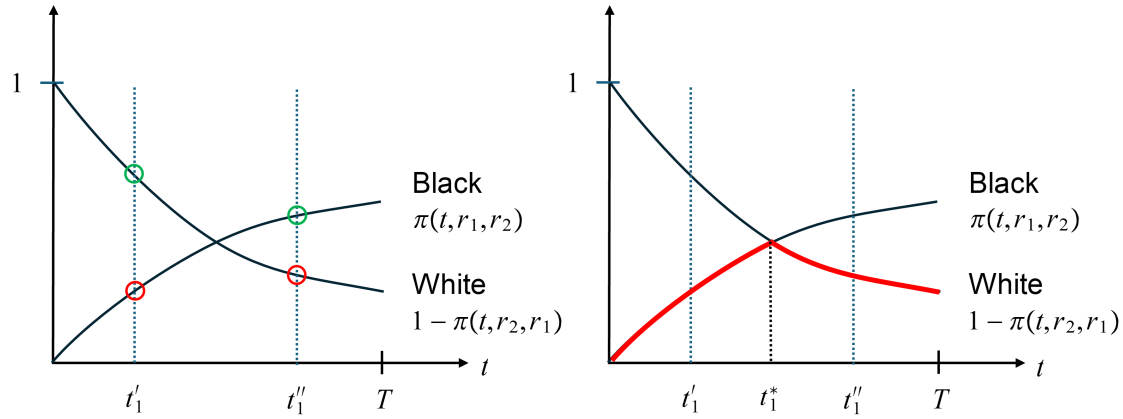


Figure 1: Player 1's winning probability with bids t_1' or t_1''

It is clear from Figure 1 that Player 1 maximizes their minimum payoff at the bid t^* that uniquely solves

$$1 - \pi(t, r_2, r_1) = \pi(t, r_1, r_2). \tag{1}$$

By symmetry, Player 2 maximizes their minimum payoff at the bid t^* that solves $1 - \pi(t, r_1, r_2) = \pi(t, r_2, r_1)$, but this equation is identical to Player 1's. Hence each player maximizes their minimum payoff with the same bid t^*, and the minimax solution is unique.

Since Armageddon bidding is a two-player zero-sum game, the minimax solution and the Nash equilibrium coincide.

Theorem: *The maxmin (Nash equilibrium) strategies are* $t_1^* = t_2^* = t^*$, *where* t^* *is the unique solution to*

$$\pi(t^*, r_1, r_2) + \pi(t^*, r_2, r_1) = 1.$$

Although it may seem counterintuitive that asymmetric players bid the same amount of time, this result follows immediately by applying Nash reasoning. Suppose the bids t_1 and t_2 are unequal and Player 1, say, has the smaller bid. Player 1's payoff is $\pi(t_1, r_1, r_2)$. Player 1's bid t_1 is not a best response to t_2 since by raising their bid to $t_1 + \varepsilon$, for $t_1 + \varepsilon < t_2$, they raise their payoff from $\pi(t_1, r_1, r_2)$ to $\pi(t_1 + \varepsilon, r_1, r_2)$.

Although Nash equilibrium coincides with the minimax solution in two-player zero-sum games, it is easy to verify directly that $t_1^* = t_2^* = t^*$, where t^* is the solution to (1), is a Nash equilibrium: Since $t_1^* = t_2^*$ the players are randomly assigned to colors.[7] Player 1 is equally likely to play Black and White, and wins with probability

$$\frac{1}{2}\pi(t^*, r_1, r_2) + \frac{1}{2}(1 - \pi(t^*, r_2, r_1)),$$

which is equal to both $\pi(t^*, r_1, r_2)$ and $1 - \pi(t^*, r_2, r_1)$ by equation (1). If Player 1 reduces their bid to $t_1' < t^*$, then Player 1 is assigned Black and wins with probability $\pi(t_1', r_1, r_2)$, which is less than their equilibrium win-

[7] Because Bidding Armageddon has a continuum of actions and discontinuous payoffs, standard Nash equilibrium existence results do not apply (see Reny (1999)). We establish existence directly by construction.

ning probability of $\pi(t^*, r_1, r_2)$ since $\pi(t, r_1, r_2)$ is increasing in t. If Player 1 increases their bid to $t_1' > t^*$, then Player 1 is assigned White and wins with probability $1 - \pi(t^*, r_2, r_1)$, which is the same as their equilibrium win probability. Thus $t_1 = t^*$ is a best response for Player 1 to $t_2 = t^*$.[8]

While we write the winning probabilities as functions of the players' ratings, this should not be taken literally. Winning probabilities depend on the identities of the two players in the match, and may vary with the time controls (e.g., bullet, blitz, or rapid) or other features of the match. As is standard in applications of game theory, we assume that the players know the payoff function π. This is plausible, given that the players in our dataset are elite international competitors who study their opponents' play.

Minimax, and equivalently Nash equilibrium, play in Bidding Armageddon yields several testable hypotheses. The first two hypotheses follow from the fact that the players' color assignments are as good as random at the equilibrium bids.

Hypothesis 1: The stronger player in a pair is equally likely to play Black or White.

Hypothesis 2: The distribution of ratings is the same for players assigned to Black and for players assigned to White.

Hypothesis 3, follows from the indifference condition that defines the equilibrium bid: a player wins with the same probability whether (i) playing Black with t^* on their clock, or (ii) playing White with T on their clock while their opponent plays Black with t^* on Black's clock. It is clear from Figure 1 that when the two bids are not exactly the same, then Hypothesis 3 approximately holds so long as the smaller of two bids is near t^*.

Hypothesis 3: A player wins with the same probability whether assigned to Black or assigned to White.

[8] Since $t_1' > t^*$ is also a best response to t^*, the Nash equilibrium is not strict.

It is easy to verify that hypotheses 1 and 3 imply that White wins with probability one half, which is Hypothesis 4.

Hypothesis 4: White wins with probability 1/2.

The final hypothesis is that the maximin (equivalently, Nash equilibrium) bids are the same.

Hypothesis 5: The players' bids are equal.

We conclude our analysis of Armageddon bidding by noting that it is a generalization of the following Divide-a-Dollar game. Consider the game in which two players simultaneously demand a share $s \in [0, 1]$ of one dollar. The player who demands the smaller share receives their demand, and the other player receives the residual. If they demand the same shares, then one player, chosen at random, receives their demand and the other player receives the residual. In the notation introduced above, this Divide-a-Dollar game is represented by the function $\pi(s, r_1, r_2) = s$. In the minimax (and Nash equilibrium) solution each player demands a one-half share and each player receives their demand. Armageddon Bidding, rather than dividing a dollar, divides a unit probability of winning a chess match.

Armageddon Bidding is fair in the sense that it treats equal players equally and it conveys an advantage to stronger players. Given equal players, i.e., $\pi(t, r_1, r_2) = \pi(t, r_2, r_1)$ for every t, then at the equilibrium bid each player wins with probability 1/2. If Player 1 is stronger than Player 2, i.e., $\pi(t, r_1, r_2) > \pi(t, r_2, r_1)$ for every t, then at the equilibrium bid Player 1 wins with probability greater than Player 2, i.e., $\pi(t^*, r_1, r_2) > 1 - \pi(t^*, r_1, r_2)$.[9]

[9]The equilibrium bid t^* satisfies $\pi(t^*, r_1, r_2) + \pi(t^*, r_2, r_1) = 1$. If $\pi(t, r_1, r_2) = \pi(t, r_2, r_1)$ for every t, then $\pi(t^*, r_1, r_2) = \pi(t^*, r_2, r_1)$ implies $\pi(t^*, r_1, r_2) = \pi(t^*, r_2, r_1) = 1/2$. If $\pi(t, r_1, r_2) > \pi(t, r_2, r_1)$ for every t, then $\pi(t^*, r_1, r_2) > \pi(t^*, r_2, r_1)$ implies $\pi(t^*, r_1, r_2) > 1/2 > 1 - \pi(t^*, r_1, r_2)$.

3 The Data

We obtained data from Chess.com on every Armageddon game it hosted between August 2022 and May 2025—a total of 523 games involving 203 players across 24 tournaments. These tournaments are elimination bracket tournaments, where in each match players play multiple games against their opponent. A player who loses a match is either eliminated from the tournament, or continues in a losers bracket. A summary of the tournaments in our data set is in Appendix Table A1. In all but one tournament, the total prize money ranged from $2,500 to $500,000, with individual prizes as large as $200,000.[10] All players in our data are titled players, most of whom (91%) are Grandmasters.[11] For each Armageddon game, we observe the players' bids, ratings, moves, time used per move, and the game outcome, tournament name, and time control. In addition, we collected all the regular games played in the same tournament bracket as an Armageddon game. This included, in particular, all the games in a match leading up to an Armageddon game.

Table 1 below presents summary statistics on outcomes for regular and Armageddon games in our data. The distribution of outcomes in the regular games is similar to historical distributions of outcomes in games played by masters. It is worth noting that all games in the tournaments in our dataset were played under short time controls, ranging from 15 minutes with no increment to 1 minute with a 1-second increment per move, depending on the tournament. Armageddon tie-break games were played with no increment. It is evident that the probability of a draw is substantially lower in Armageddon

[10] In the Olympic Esports Series 2023 tournament, the top 4 players qualified for a final-four tournament organized by the International Olympic Committee with no prize money.

[11] Entry to one tournament was restricted to titled players who had not attained the Grandmaster title.

games than in regular games, while the probability of a White win is higher.

	Regular Games	Armageddon
White Win	36.2%	49.7%
Draw	36.3%	21.6%
Black Win	27.5%	28.7%
Avg. Rating	2705.00	2667.62
	(174.60)	(105.40)
Unique players	273	203
Number of games	5,252	523
Number of matches	1,775	523

Standard errors in parentheses.

Table 1: Summary Statistics

The majority of matches consisted of three games (364 matches or 69.6%), followed by five-game matches (147 matches or 28.1%). Matches of seven or more games were rare, accounting for only 2.3% of the sample.[12] In the Armageddon games, 426 of the 523 pairs (81.5%) of players faced each other only once; 33 pairs faced each other twice.

4 Results

We distinguish the players in a match by their ratings, higher (stronger) or lower (weaker). According to Hypothesis 1, the stronger player is equally likely to play Black or White. We find that the stronger player in each match plays White in 49.8% of the games, which is statistically indistinguishable

[12] In bullet-format tournaments (in which each player starts with one minute on the clock), players compete for a fixed period, usually 15 minutes, rather than playing a fixed number of games. In our data, the number of regular games was as high as 40.

from the theoretical prediction of 50% ($p = 0.965$, $N = 520$ binomial test).[13] Because ratings are a noisy measure of player strength, we verify the robustness of this result by imposing increasingly stringent definitions of "stronger." Specifically, we classify player i as stronger than player j only when i's rating advantage, $d_{ij} = r_i - r_j$, exceeds some minimum threshold, dropping all pairs below this cutoff. This approach trades statistical power for a cleaner signal of relative strength.

Figure 2 plots the frequency with which the stronger player is White for minimum ratings differences ranging from 1 point (yielding 520 games) to 100 points (yielding 118 games). Across the entire range, the observed frequency remains tightly clustered around 50%. The null hypothesis that the stronger player plays White is not rejected at a 5% significance level for any threshold d_{ij}.

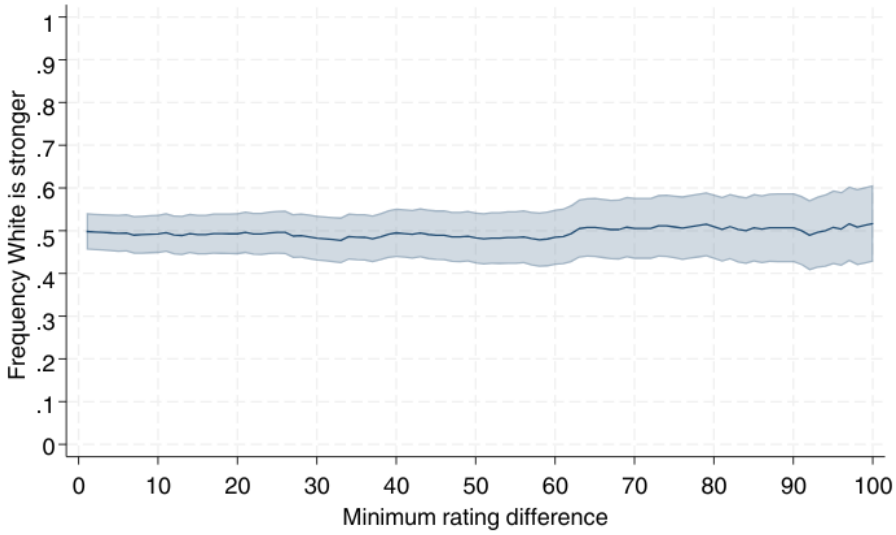


Figure 2: Frequency the stronger player is White

The consistency of this result across varying definitions of strength provides

[13] In all tests of hypotheses relating to the stronger player we drop the three Armageddon games in which both players had the same rating.

strong support that color assignment is indeed independent of relative player ability, consistent with Hypothesis 1.

According to Hypothesis 2, rating distributions do not differ by player color. Figure 3 shows the empirical *c.d.f.*'s of player ratings for Black and White. The null hypothesis that the ratings of Black and White players are drawn from the same distribution is not rejected by the two-sided Kolmogorov-Smirnov test ($D = 0.055$, $p = 0.397$). The average ratings for Black (2667.39) and White (2667.85) are virtually identical, consistent with Hypothesis 2.[14]

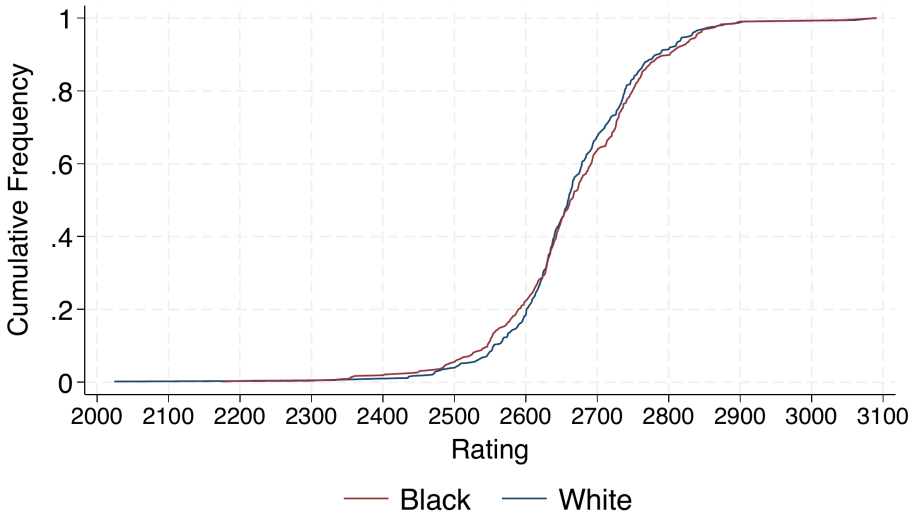


Figure 3: Empirical *c.d.f.* of player ratings, by color

A key feature of Armageddon bidding is that a player is theoretically neither advantaged nor disadvantaged by their color assignment. In other words, a player wins with the same probability whether assigned Black or White. There are 520 games in which there is a stronger player. Of these, in 261 games the stronger player was assigned Black and won 138 of those games, for a win rate of 52.9%. In 259 games, the stronger player was assigned

[14] The averages are not statistically different ($t(522) = 0.11$, $p = 0.912$).

White and won 135 of those games, for a win rate of 52.1%. We cannot reject the null hypothesis that the stronger player wins with the same probability whether assigned to Black or White ($z = 0.171$, $p = 0.864$, two-sample proportions test).[15]

We verify the robustness of this result by comparing the win rates of the stronger player, when Black and when White, for increasingly stringent definitions of stronger. Figure 4 shows the win rate of the stronger player for each color, as a function of the minimum rating difference. Not surprisingly, the win rates of the stronger player tend to increase in the rating difference, whether playing Black or White. The win rates are not, however, significantly different for any rating difference threshold.

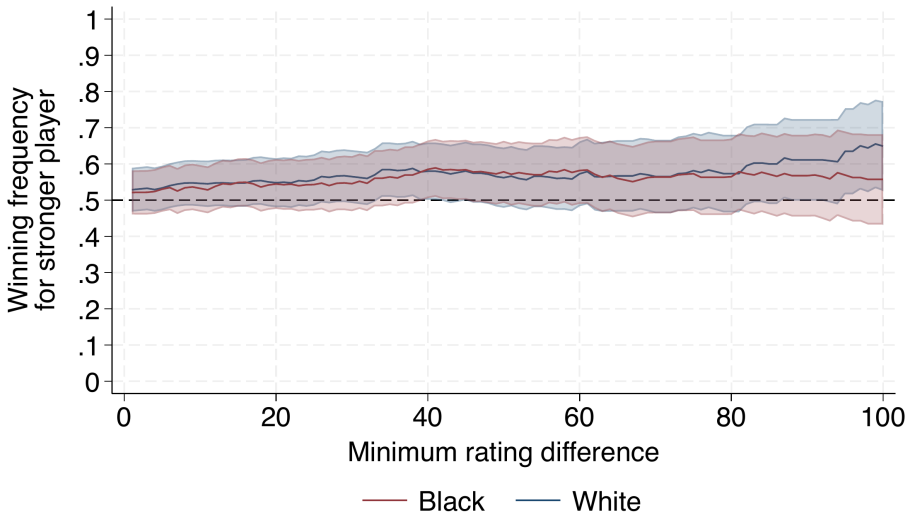


Figure 4: Win rate of the stronger player, by color

We have shown, consistent with Hypothesis 3, that stronger players win Armageddon games at the same rate with either color. To further substantiate this result, Table 2 presents OLS regressions in which the dependent variable is 1 if the stronger player won the Armageddon game and is 0 oth-

[15] In 13 games, the winner chose White. Excluding these games from the analysis does not alter our results. See Table B2 in the Appendix.

erwise.[16] We estimate the model

$$\mathbf{1}_k^{Win} = \beta_0 + \beta_1 \times \mathbf{1}_k^{White} + \beta_2 \times BlackBid_k + \beta_3 \times \mathbf{1}_k^{White} \times BlackBid + \beta_4 \times EPS_{ik} + \gamma X + \varepsilon_k,$$

where $\mathbf{1}_k^{Win} \in \{0, 1\}$ is an indicator which takes the value 1 if in the k-th Armageddon game the stronger player won, $\mathbf{1}_k^{White}$ is an indicator which takes the value 1 if the stronger player played White, $BlackBid_k$ is the number of seconds on Black's clock at the start of the game, EPS_{ik} is the Elo Predicted Score of player i (the stronger player) given their own and their opponent's ratings, and X is a vector of controls for player identity, game time, and Fischer Random rules.[17,18] The EPS of player i, with rating r_i, in a game against player j, with rating r_j, is given by the formula

$$EPS(r_i, r_j) = \frac{1}{1 + 10^{(r_j - r_i)/400}}.$$

$EPS(r_i, r_j)$ is interpreted as the expected number of points that player i wins in a regular game of chess against player j, where a win counts for 1, a draw counts for .5, and a loss counts for 0. In Armageddon games, a draw is not a possible outcome and $EPS(r_i, r_j)$ can be interpreted as the probability

[16] We report linear probability model estimates for ease of interpretation. Logit estimates with player controls drop 78 observations for those players that won all of their Armageddon games. However, the qualitative results remain the same.

[17] One tournament in our data was played under Fischer Random rules, in which the pieces in the back rank are randomly placed at the start of the game.

[18] The time on Black's clock at the start of the game is Black's bid, except for the 13 games in which the low bidder chose White. In these games, Black plays with a full clock and $BlackBid$ is set to the full clock.

that i wins against j, given their rating difference.[19]

Dependent Variable: Stronger Player Wins					
	(1)	(2)	(3)	(4)	(5)
White	−0.003	0.002	0.005	0.117	0.292
	(0.052)	(0.053)	(0.052)	(0.174)	(0.231)
EPS		0.937**	0.941**	0.788***	0.920**
		(0.412)	(0.412)	(0.283)	(0.419)
Black Bid	0.000		0.000	0.000	0.001
	(0.001)		(0.001)	(0.000)	(0.001)
White×Black Bid				−0.000	−0.001
				(0.000)	(0.000)
Player Controls	Yes	Yes	Yes	No	Yes
Time Format Controls	Yes	Yes	Yes	Yes	Yes
Fischer Random Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.253	0.267	0.267	0.022	0.271
Observations	520	520	520	520	520

Standard errors (in parentheses) are clustered at the player level; *** p<0.01, ** p<0.05, *p<0.10.

Table 2: Determinants of Stronger Player Winning in Armageddon

The results in Table 2 show, for every specification, that neither the strong player's assigned color nor the time on Black's clock at the start of the game has a statistically significant effect on the stronger player's likelihood of winning. EPS_{ik}, by contrast, is consistently highly statistically significant. In our preferred specification (5), there is a nearly one-to-one correspondence between the stronger player's ex-ante win probability, as captured by EPS, and actual outcomes: for a one percentage point increase in the stronger player's EPS expected winning probability, the stronger player's probability

[19] A player typically has separate ratings for different time controls, such as classical, rapid, blitz, and bullet. In the regression, we use the rating corresponding to the tournament's time control.

of winning increases by 0.920 percentage points. We cannot reject the null hypothesis that $H_0 : \beta_4 = 1$, i.e., the coefficient on EPS_{ik} is 1 ($F(1, 130) = 0.04$ and $p = 0.850$). The message from these results, across all specifications, is consistent: the Armageddon tiebreak format eliminates any systematic advantage from playing White or Black, and it preserves the skill hierarchy implied by the players' Elo ratings.

According to Hypothesis 4, White wins with probability .5. In the data, White won 49.7% of the Armageddon games. The null hypothesis that White wins with probability .5 is not rejected ($p = 0.930$, $N = 523$ binomial test).

Bidding Behavior

Although theory predicts that the two players make identical bids, there were no games in our data in which the two bids were exactly equal. However, that prediction is predicated on the assumption that both players know the payoff functions $\pi(t, r_1, r_2)$ and $\pi(t, r_2, r_1)$ perfectly, and both act, without noise or error, to maximize their minimum payoff. To evaluate the conformity of the bids to theory, we compare the bids of the stronger and weaker players. Figure 5 shows a scatter plot of 520 bid pairs, where the time bid by the weaker player is measured on the horizontal axis and the time bid by the stronger player is measured on the vertical axis. Colors reflect the time controls: bids in 5 minute games (300 seconds) are shown in blue, in 10 minute games are shown in red, and in 15 minute games are shown in green. It is evident that the bid pairs are centered around the 45-degree line, with most close to the line, which suggests that stronger players do not systematically bid more aggressively than weaker players (or vice versa).[20] The average bids are virtually the same: stronger players bid an average 0.76 second less than weaker players, a difference which is not statistically different from zero

[20] The Pearson's correlation between the bids of the stronger and weaker player is 0.750 and is statistically significant ($p < 0.0001$).

($t(519) = -0.17$ and $p = 0.861$).

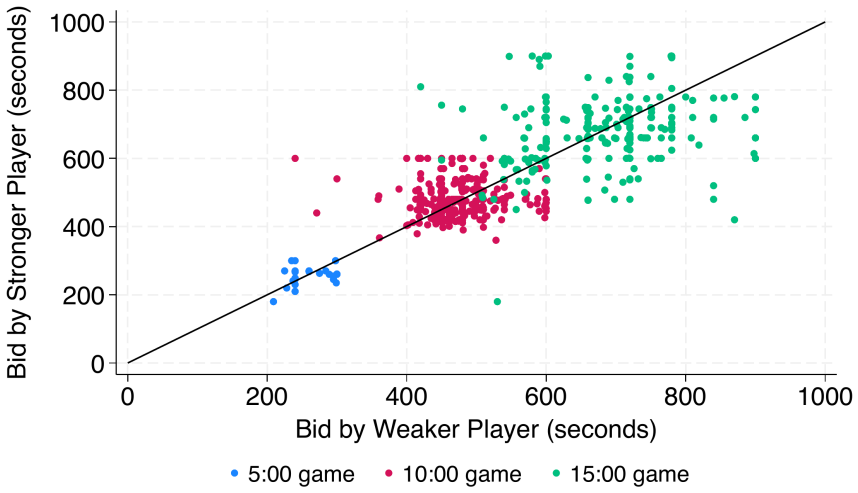


Figure 5(a): Bids by Stronger and Weaker Player

Figure 5(b) shows a histogram of bid differences, White bid minus Black bid, as a percentage of the full clock. The majority of the bids were close: in 62% of all games, the bid difference is less than 10 percent of the full clock time. The histogram to the left of zero shows the 13 observations in which

the player with the lower bid chose White.[21]

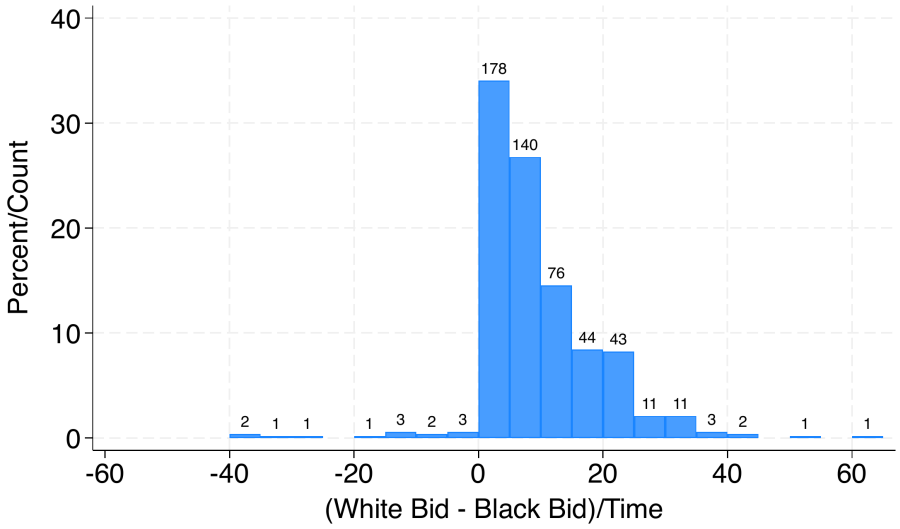


Figure 5(b): Histogram of White minus Black bids

Table 3 below presents OLS regressions analyzing the determinants of bids. We estimate the model

$$b_{ik} = \beta_0 + \beta_1 \times \%BlackWinsLeadUp_k + \beta_2 \times EPS_{ik} + \gamma X + \varepsilon_{ik},$$

where the dependent variable b_{ik} is time bid as a proportion of the full clock time T by player i in Armageddon game k, EPS_{ik} is the Elo predicted score for player i in game k, and X is a vector of controls for player, game time, and Fischer Random rules. On average, players bid approximately 78% of T in order to play Black. The variable *% Black Wins Lead Up* is the percentage of regular games in a match won by Black. For instance, if Black won all four lead-up games to the tiebreak, then its value would be 1.

[21] We believe that these observations were errors, resulting from inexperience with the format. 12 of the 13 observation occur in the first half of the sample. It was the first Armageddon game for 8 of these players, and all but one player made this decision only once. White won only 3 of these 13 games.

If Black won two games and the other two games were draws, then its value would be 0.5. The results show that the coefficient on *% Black Wins Lead Up* is negative and statistically significant. In our preferred specification (3), players bid approximately 2.0 percentage points lower when Black won all the lead-up games than when Black lost them all. In the context of a 15-minute Armageddon game, this translates into bids that are, on average, 18 seconds shorter, reflecting more aggressive bidding in favor of playing Black.[22,23]

Dependent Variable: Bid as a % of T			
	(1)	(2)	(3)
% Black Wins Lead Up	-0.020^{**}		-0.020^{**}
	(0.010)		(0.010)
EPS		0.020	0.022
		(0.028)	(0.028)
Constant	0.787^{***}	0.773^{***}	0.776^{***}
	(0.002)	(0.014)	(0.014)
Player Controls	Yes	Yes	Yes
Time Format Controls	Yes	Yes	Yes
Fischer Random Controls	Yes	Yes	Yes
R-squared	0.560	0.557	0.560
Observations	1046	1046	1046

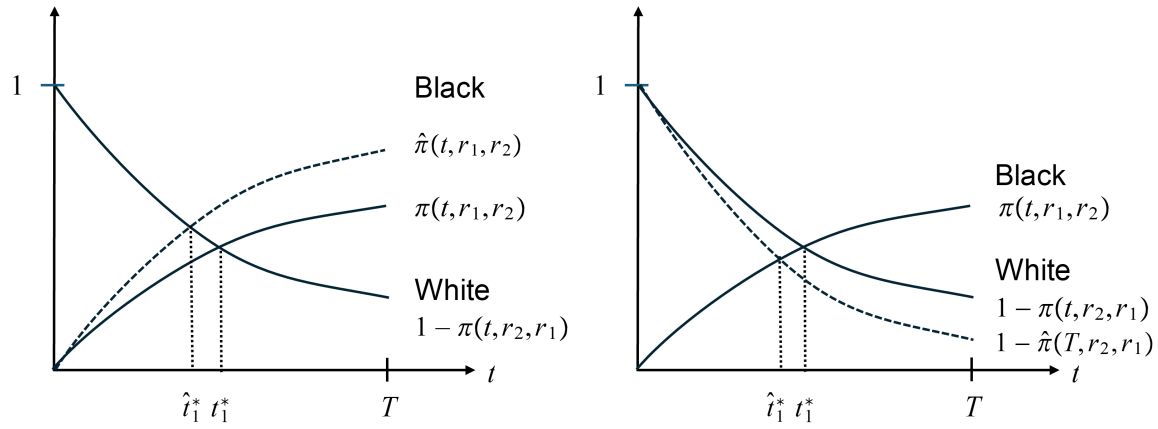
Standard errors (in parentheses) are clustered at the player level; *** p<0.01, ** p<0.05, *p<0.10.

Table 3: Determinants of Bids

[22] Across all specifications, the relative strength of the players, as captured by EPS, has no statistically significant effect on bidding behavior, whether we consider a linear or quadratic specification.

[23] Standard errors can be clustered at either the game or player level. Game-level clustering permits arbitrary correlation between the two bids within a Bidding Armageddon game, but assumes independence between games. Player-level clustering permits correlation between bids made by the same player in different games, but assumes independence between the two players' bids within a game. Results are virtually identical under either approach.

The negative coefficient estimate on *% Black Wins Lead Up* is consistent with the theoretical model, which predicts that when players are stronger at playing Black, then the equilibrium bid is smaller. As shown in Figure 6(a), if Player 1 is stronger at playing Black, holding their ability as White fixed, then their payoff function $\pi(t, r_1, r_2)$ shifts upward to $\hat{\pi}(t, r_1, r_2)$. Likewise, if Player 2 is stronger at black, then $1-\pi(t, r_2, r_1)$ shifts down to $1-\hat{\pi}(t, r_2, r_1)$. In either case, the equilibrium bid declines. Hence, in matches in which one or both players are especially strong as Black, we expect the win rate for Black in the lead-up games to be higher and the equilibrium bids in the Armageddon tiebreak to be lower. This is precisely what Table 3 reveals.



(a) Player 1 Stronger at Black (b) Player 2 Stronger at Black

Figure 6: Player 1's winning probability

5 The Mechanism

The preceding results demonstrate that Armageddon bidding successfully neutralizes color advantage: stronger players win at the same rate regardless of their color assignment. This section investigates the mechanism by which this result is achieved. We show that White's time advantage translates into more accurate play. This increase in accuracy exactly offsets Black's

advantage from draw odds, thereby equalizing winning probabilities between colors.

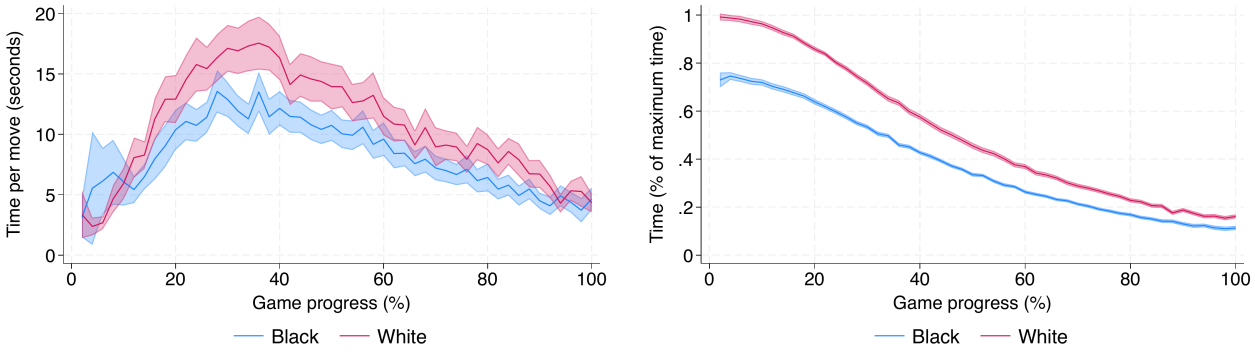
Time Use

Players in timed chess face a dynamic resource allocation problem: time spent analyzing each move reduces the risk of making an error but increases the risk of losing on time. Figure 7(a) shows how Black and White allocate time in Armageddon chess. The horizontal axis shows the move number, expressed as a proportion of the total number of moves made by the player, thereby allowing comparisons across games of different lengths.[24] The vertical axis shows the number of seconds. The figure shows that there is no significant difference in time use between Black and White early in a game. Both players move quickly at the opening. As the game progresses, and the board becomes more complex, both players spend more time per move. White, however, spends significantly more time than Black until near the end of the game. This suggests that White calculates more deeply, evaluates more alternatives, and verifies candidate moves more carefully.

Figure 7(b) shows time remaining, as a percentage of total time T, as a function of game progress. Black begins with approximately 78% of White's time. It shows that White's initial time advantage over Black shrinks, but

[24] The figure shows time use starting from the third move by White. Black's start time was manually entered by Chess.com staff, and hence for the first couple of moves time used by Black was not always accurately reflected in the Portable Game Notation (PGN) files.

never entirely vanishes.



(a) Time used per move

(b) Time remaining

Shaded areas are 95% confidence intervals based on clustered standard errors at the game level.

Figure 7: Time use

Accuracy of Play

Next we evaluate how time use translates into quality of play. To quantify the accuracy of play, we employ Winter, an AI chess engine with native support for Armageddon chess. Given the history of play at the t-th move, Winter assigns probabilities p_t^W, p_t^D, and p_t^B that the final outcome of the game is, respectively, a win for White, a draw, or a win for Black, under computer-optimal play by both players.[25] We use this probability distribution as a benchmark for optimal play, subject to several caveats. First, different chess engines produce slightly different evaluations of the same position. Second, Winter assumes "perfect" play by both players. That is, when evaluating the board at White's move, Winter assumes that at subsequent moves White will act to maximize its probability of winning, while Black will act to minimize White's winning probability. In other words, it assumes minimax play. Third, the depth of the engine's analysis is constrained by

[25] We use Winter to evaluate each board position to a depth of 30. Engine depth measures how many moves ahead the engine calculates. Depth 30 represents superhuman analysis, typically evaluating millions of positions per move.

available computational resources. Finally, the engine's evaluations do not account for the fact that players are subject to time constraints and that identifying strong moves is time-consuming.

Figure 8 displays p_t^W, p_t^D, and p_t^B, averaged across all games, as play progresses. At the start of an Armageddon game, White wins with probability 0.233, the outcome is a draw with probability 0.573, and Black wins outright with probability 0.193. These probabilities remain relatively stable through the first half of play, but then diverge sharply as time pressure mounts and errors accumulate. Naturally, Winter's board evaluations near the end of the game closely match the actual outcome frequencies (see Table 1). It is evident from the figure that human play does not proceed as Winter predicts, since Winter's evaluations change over time. Notably, Winter's evaluation of White's probability of winning approximately doubles from when the game starts to when it ends. This suggests that White's play is superior to Black's play.

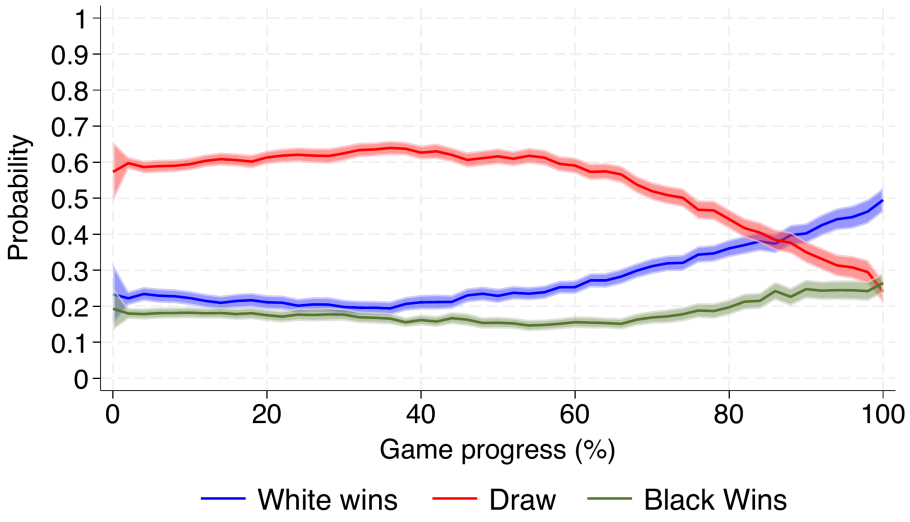


Figure 8: Winter predicted game outcome

From Winter's evaluations we construct two outcome variables. We define

the *Accuracy* of White's move at t as the difference

$$p_{t+1}^W - p_t^W,$$

i.e., it is White's winning probability immediately following White's move minus White's winning probability immediately prior to White's move.[26] In the data, the mean accuracy of White is -0.017, meaning White loses 1.7 percentage points of winning probability for each move relative to what Winter identifies as White's optimal move.[27]

The *Accuracy* of Black's move at t is the difference

$$p_{t+1}^D + p_{t+1}^B - (p_t^D + p_t^B),$$

i.e., it is Black's winning probability immediately following Black's move minus Black's winning probability immediately prior to Black's move. The mean accuracy of Black is -0.022. Human moves typically fall short of being computer optimal as a result of cognitive limitations and time pressure.

We define $\Delta Draw$ of the move at t as the difference

$$p_{t+1}^D - p_t^D,$$

i.e., it is the increase in the probability that Winter assigns to a draw as a result of the move at t. $\Delta Draw$ is useful to understand the extent to which players' moves are shaped by the strategic incentives of Armageddon rules. We expect White to play to reduce $\Delta Draw$, whereas Black plays to either win outright or win on a draw.

[26] Since White moves first, White has the move when t is odd and Black has the move when t is even.

[27] By Zermelo's Theorem, at any board position in Armageddon chess one of the players has a winning strategy, i.e., a strategy that guarantees that the player wins. If White has a winning strategy, then a perfect chess engine would assign accuracy 0 to an optimal move since White has a winning strategy both before and after his move. It would assign accuracy -1 to a suboptimal move since White had a winning strategy before their move, but Black had a winning strategy after White's move. If the solution to chess were known, then accuracy would be simple, but uninteresting.

Table 4 presents regression results on the determinants of play. We estimate the model

$$Y_{itk} = \beta_0 + \beta_1 \times \mathbf{1}_{ik}^{White} + \beta_2 \times Time_{itk} + \beta_2 \times \mathbf{1}_{ik}^{White} \times Time_{itk} + \gamma_k + \varepsilon_{itk},$$

where Y_{itk} is either *Accuracy* or $\Delta Draw$. $Time_{itk}$ is the time on player i's clock at their t-th move in game k. $\mathbf{1}_{ik}^{White}$ is an indicator function which takes the value 1 if player i is White in game k, and γ_k are game fixed effects. The dependent variable in columns (1) through (3) is *Accuracy*. The estimated constant in (1) is negative, which confirms that humans consistently fall short of computer-optimal play, losing on average 2.245 percentage points of winning probability per move. The specification in (2) includes game controls to account for any game-specific heterogeneity. It reveals the central finding: White loses 0.486 fewer percentage points of winning probability per move than Black. The greater accuracy of White directly reflects the greater time spent per move by White.

Column (3) directly estimates the marginal value of time and shows that time is valuable for both players, but more valuable for Black. Each additional second on Black's clock increases the accuracy of Black's move (and winning probability) by 0.005 percentage points. Each additional second on White's clock increases the accuracy of White's move by 0.002 percentage

points.[28]

Dep. Var.	Accuracy			ΔDraw	
	(1)	(2)	(3)	(4)	(5)
White	0.482***	0.486***	1.104***	−1.748***	−1.748***
	(0.042)	(0.042)	(0.086)	(0.379)	(0.381)
Time			0.005***		
			(0.001)		
White×Time			−0.003***		
			(0.000)		
Constant	−2.245***	−2.247***	−3.634***	0.539**	0.539**
	(0.118)	(0.021)	(0.154)	(0.192)	(0.190)
Game Fixed Effects	No	Yes	Yes	No	Yes
R-squared	0.000	0.028	0.030	0.002	0.003
Observations	56,764	56,764	56,764	56,764	56,764

Standard errors (in parentheses) are clustered at the player level; *** p<0.01, ** p<0.05, *p<0.10.

Table 4: Determinants of Play

Under Armageddon rules the players have strictly opposing incentives for a draw: White actively avoids draws since draws award victory to Black. Black, conversely, is indifferent between drawing and winning, as both yield victory. Column (4) of Table 4 tests this prediction by examining how moves by Black and White change the probability of a draw. It shows that each move by Black increases the probability of a draw by 0.539 percentage points. By contrast, a move by White reduces the probability of a draw by 1.748 percentage points (relative to a move by Black). The coefficient estimates remain the same when we add game controls.

The results of this section show that Armageddon bidding converts time into playing strength: by accepting less clock time, Black obtains draw odds,

[28] The effect for White, Time+White×Time = .002, is statistically significantly different from zero, $F(1, 521) = 32.82$, $p < .001$.

but at the cost of reduced accuracy. The trade-off between time and accuracy exactly balances, as theory predicts, since a player wins with the same probability whether Black or White. Further, Armageddon rules influence the players' strategies. Black tends to steer the games towards a draw, while White plays to avoid a draw.

6 Conclusion

We provide a first test using field data of the Minimax Hypothesis and the theory of pure strategy Nash equilibrium. Armageddon bidding is an ideal setting to cleanly test theory for several reasons: First, the players are the world's elite professional chess players, who know the payoff function of the game, and who play for substantial monetary stakes. Second, the game has only two outcomes, and therefore equilibrium does not depend on risk attitudes. Third, the game is strictly competitive, so there is no scope for cooperation or social preferences to shape behavior. We find that game outcomes are consistent with theory and, while bids are not equal, they are close. In particular, a player's probability of winning the Armageddon game is the same regardless of whether they play Black or White.

Armageddon bidding is a novel game format and the payoffs of the game differ from one pair of players to the next. Hence the conformity of play to equilibrium that is observed in the data cannot be the result of learning through repeated play. This suggests that equilibrium play is the result of minimax reasoning – players choose their bids to maximize their minimum payoff. In doing so, players display a deep understanding of the trade-off between clock time and accuracy of play.

References

[1] Anderson, Axel, Rosen, Jeremy, Rust, John, and Kin-Ping Wong. (2025). "Disequilibrium Play in Tennis." *Journal of Political Economy,*

133, 190-251.

[2] Brown, James and Robert Rosenthal (1990), "Testing the Minimax Hypothesis: A Reexamination of O'Neill's Experiment." *Econometrica*, 58, 1065-1081.

[3] Chiappori, Pierre, Levitt, Steven, and Timothy Groseclose (2002), "Testing Mixed Strategy Equilibria When Players are Heterogeneous: The Case of Penalty Kicks in Soccer." *American Economic Review*, 92, 1138-1151.

[4] Costa-Gomes, Miguel and Georg Weizsäcker (2008), "Stated Beliefs and Play in Normal-Form Games." *The Review of Economic Studies*, 75, 729–762.

[5] Gauriot, Romain, Page, Lionel, and John Wooders (2023), "Expertise, Gender, and Equilibrium Play." *Quantitative Economics*,14, 981-1020.

[6] Gonzalez-Diaz, Juan, Gossner, Olivier, and Brian Rogers (2012), "Performing Best When It Matters Most: Evidence From Professional Tennis." *Journal of Economic Behavior & Organization*, 84, 767-781.

[7] Hsu, Shih-Hsun, Huang, Chen-Ying and Cheng-Tao Tang (2007), "Minimax Play at Wimbledon: Comment." *American Economic Review*, 97, 517-523.

[8] Levitt, Steven, List, John, and David Reiley (2010), "What Happens in the Field Stays in the Field: Professionals Do Not Play Minimax in Laboratory Experiments." *Econometrica*, 78, 1413-1434.

[9] Levitt, Steven, List, John, and Sally Sadoff (2011), "Checkmate: Exploring Backward Induction among Chess Players," *American Economic Review*, 101, 975-990.

[10] Nash, John (1950), "Equilibrium Points in N-Person Games." *Proceedings of the National Academy of Sciences*, 36, 48-49.

[11] Nash, John (1953), "Two-Person Cooperative Games." *Econometrica*, 21, 128–140.

[12] Nagel, Rosemary (1995). "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*, 85, 1313–1326.

[13] O'Neill, Barry (1987), "Nonmetric Test of the Minimax Theory of Two-Person Zero-Sum Games." *Proceedings of the National Academy of Sciences*, 84, 2106-2109.

[14] O'Neill, Barry (1991), "Comments on Brown and Rosenthal's Reexamination." *Econometrica*, 59, 503-507.

[15] Palacios-Huerta, Ignacio (2003), "Professionals Play Minimax." *Review of Economic Studies*, 70, 395-415.

[16] Reny, Philip (1999), "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discontinuous Games," *Econometrica*, 67, 1029–1056.

[17] Paserman, Daniele (2023), "Gender Differences in Performance in Competitive Environments? Evidence from Professional Tennis Players." *Journal of Economic Behavior & Organization*, 212, 590-609.

[18] Roth, Alvin (1993). "The Early History of Experimental Economics." *Journal of the History of Economic Thought*, 15, 184-209.

[19] Salant, Yuval and Jorg Spenkuch (2026), "Complexity and Satisficing: Theory with Evidence from Chess," *Review of Economic Studies* 93, 1296-1322.

[20] Shachat, Jason (2002), "Mixed Strategy Play and the Minimax Hypothesis." *Journal of Economic Theory*, 104, 189-226.

[21] Stahl, Dale and Paul Wilson (1994), "Experimental evidence on players' models of other players." *Journal of Economic Behavior & Organization*, 25, 309–327.

[22] Stahl, Dale and Paul Wilson (1995). "On players' models of other players: Theory and experimental evidence." *Games and Economic Behavior*, 10, 218–254.

[23] Van Essen, Matthew and John Wooders (2015), "Blind Stealing: Experience and Expertise in a Mixed-Strategy Poker Experiment." *Games and Economic Behavior*, 91, 186-206.

[24] Von Neumann, John (1928). "Zur Theorie der Gesellschaftsspiele." Math. Ann. 100: 295–320.

[25] Walker, Mark and John Wooders (2001), "Minimax Play at Wimbledon." *American Economic Review*, 91, 1521-1538.

[26] Wooders, John (2010), "Does Experience Teach? Professionals and Minimax Play in the Lab." *Econometrica*, 78, 1143–1154.

7 Appendix

Tournaments

Year	Tournament	Total Prize Money	Individual Prize Range	Other Prize
2022	Chess.com Global Championship Finals	$500,000	$25,000-$200,000	-
2022	Chess.com Global Championship Knockout	$440,000	$5,000-$15,000	Spot in Finals for Winner
2022	Chess.com Rapid Chess Championship Finals	$150,000	$4,000-$30,000	-
2022	Fischer Random WC Qualifier	$2,500	$200-$500	Spot in Finals for Winner
2022	Speed Chess Championship Qualifiers	$10,000	$400-$1,000	Spot in Main Event for Winner
2023	Champions Chess Tour, AI Cup	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour, Aimchess Rapid	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour, Airthings Masters	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour, Chessable Masters	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour, ChessKid Cup	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour, Julius Baer Generation Cup	$235,000	$1,000-$30,000	CCT tour points
2023	Champions Chess Tour Finals	$500,000	$1,000-$200,000	**
2023	I'M Not A GM Speed Chess Championship	$15,000	*	-
2023	Junior Speed Chess Championship Qualifier	$2,500	$150-$500	Spot in Finals for top 2
2023	Olympic Esports Series	-	-	Spot in Finals for top-4
2023	Speed Chess Championship	$150,000	$3,000-$30,000	By win %
2023	Julius Baer Women's Speed Chess Championship Qualifier	$5,000	$150-$500	-
2024	Champions Chess Tour, Chessable Masters	$300,000	$1,000-$30,000	CCT tour points
2024	Champions Chess Tour, Chess.com Classic	$300,000	$1,000-$30,000	CCT tour points
2024	Champions Chess Tour, Crunchlabs Masters	$287,500	$1,500-$30,000	CCT tour points
2024	Champions Chess Tour, Julius Baer Generation Cup	$287,500	$1,500-$30,000	CCT tour points
2024	Champions Chess Tour Finals	$500,000	$1,000-$200,000	**
2025	Champions Chess Tour, Chessable Masters	$150,000	$5,000-$25,000	CCT tour points
2025	Champions Chess Tour, Chess.com Classic	$150,000	$5,000-$25,000	CCT tour points

Note: *: The prize structure for this tournament involved a $200 per win and $50 per loss in the group stage. In the semi-finals, the winner received $1,250, and both players split another $1,250 based on win %. **: $1,000 per win in round-robin stage.

Table A1: Tournaments

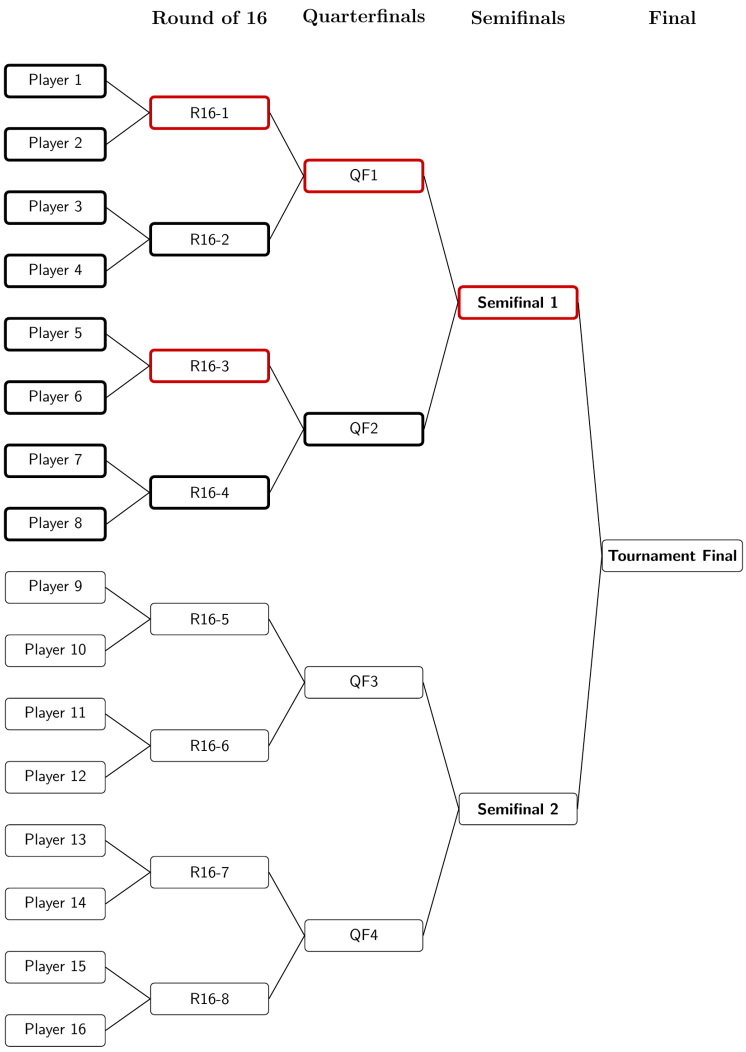


Figure A1: Stylized example of an elimination tournament, with matches resolved by an Armageddon game shown with a red border. We collected results from all matches with thick borders, but not from matches with a thin border.

Robustness Check

The hypothesis tests in the paper are based on the full sample of 523 Armageddon games. Table B2 reports the key summary statistics of the full sample to the restricted sample of 510 Armageddon games that results from dropping the 13 games in which the smaller bidder chose to play White. The sample size for tests that involve conditioning on the stronger player's outcome is 507; this is because in three Armageddon games, players had the same rating.

	White is	Average Rating		Win Freq. Stronger Player	
	Stronger	White	Black	When White	When Black
Full	49.8%	2667.386	2667.853	52.1%	52.9%
($N = 523$)		(101.206)	(109.527)		
Restricted	49.5%	2669.373	2670.661	53.4%	52.7%
($N = 510$)		(101.108)	(106.989)		

Table B2: Summary Statistics for Restricted Sample

For the restricted sample, we find that the stronger player in each match plays White in 49.5% of the games, which is statistically indistinguishable from the theoretical prediction of 50% ($p = 0.859$, $N = 507$ binomial test). Nor can we reject the hypothesis that the average rating of Black and White is the same ($t(509) = 0.31$, $p = 0.755$). Finally, the null hypothesis that the ratings of Black and White players are drawn from the same distribution is not rejected by the two-sided Kolmogorov-Smirnov test ($D = 0.055$, $p = 0.426$). Finally, we cannot reject the null hypothesis that the stronger player wins with the same probability when Black or White ($z = 0.147$, $p = 0.883$). Thus the key features of the data are the same in both the full and restricted sample.