Complexity and Choice*

Yuval Salant, *Northwestern University*
Jörg L. Spenkuch, *Northwestern University*

July 2022

Abstract

We develop a model of satisficing with evaluation errors that incorporates complexity at the level of individual alternatives. In addition to making sharp predictions about the effect of complexity on choice probabilities, the theory yields a new empirical test that leverages complexity to distinguish satisficing from a large class of maximization-based choice procedures. We test and confirm the model predictions in a novel data set with information on hundreds of millions of chess moves by highly experienced players. We further document that skill and time moderate the adverse effect of complexity on the quality of decision-making, and that they complement each other in doing so. Our findings help to shed some of the first light on how complexity affects choice behavior outside of the laboratory.

*For helpful suggestions we thank Sandeep Baliga, Tim Feddersen, Peter Klibanoff, Daniel Martin, Pablo Montagnes, and GM David Smerdon, as well as audiences at Boston College, Monash University, Northwestern, Tel-Aviv University, University of Alicante, UC Santa Barbara, the 2022 Barcelona Summer Forum, 2022 North American Summer Meetings of the Econometric Society, and the 2021 SITE Conference. This research was supported in part through the computational resources provided by the Quest high-performance computing facility at Northwestern University. Correspondence can be addressed to the authors at: MEDS Department, Kellogg School of Management, 2211 Campus Dr, Evanston, IL 60208; or by email to `y-salant@kellogg.northwestern.edu` and `j-spenkuch@kellogg.northwestern.edu`.

1. Introduction

In many choice situations, the available alternatives are inherently complex. Insurance contracts, for instance, might consist of tens or even hundreds of clauses that jointly determine value. Durable goods can have dozens of relevant attributes, and policy proposals often contain a great number of details that bedevil decision-making. Strategies in dynamic games sometimes include so many contingencies that identifying an optimal one may become nearly impossible. The goal of this paper is to better understand how complexity of this kind affects choice behavior.

Drawing on the notion of Kolmogorov complexity in information theory, we conceive of an object's complexity as the length of its efficient representation. For example, in a dynamic game, the inherent complexity of an action may correspond to the size of the ensuing subgame. The complexity of a contract might be approximated by the number of non-redundant clauses; and in the context of durable goods, complexity may correspond to the length of the array of relevant attributes.[1] We propose that alternatives with longer efficient representations are more challenging to evaluate, and that this difficulty manifests itself in decision-making as errors in the perception of value.

Our analysis begins by incorporating this idea into an empirically testable theory of choice. The theory has two main ingredients. The first one adapts the familiar random-utility framework to study complexity. Each alternative is characterized by its value to the decision maker (DM) and its inherent complexity, i.e., the length of its efficient representation. When assessing an alternative's value, the DM obtains an unbiased estimate whose accuracy depends negatively on the object's complexity.

The second ingredient of our theory incorporates this noisy evaluation process into a model of choice. Here, we build on Simon's (1955; 1972) seminal work on bounded rationality and satisficing. According to Simon, individuals may not consider all possible options and pick the best, but examine a rather small number, making a choice as soon as they find one alternative that they regard as satisfactory. In our model, the decision maker has in mind an aspiration level that she wishes to exceed. She lists all available alternatives in random order, and sequentially evaluates them until she encounters one whose estimated value exceeds the aspiration level. This is the alternative she chooses.

After developing the model's predictions, we test them in a novel dataset on chess endgames. Chess is an environment that admits an unambiguous measure of value. Every board configu-

[1] For longer object representations to imply higher complexity, all alternatives have to be described in the same language, although the language need not be the same across contexts. Unlike Kolmogorov complexity, we do not insist on the shortest possible representation. In fact, in some economic environments, strategic parties, such as firms, can deliberately simplify or complicate object descriptions in order to achieve their goals.

ration in a chess game corresponds to a choice set in which the alternatives are all available legal moves. By Zermelo's Theorem (1913), any chess move is of one of three types. A winning move allows the current player to move to force a win. A losing move enables her opponent to guarantee himself a win, whereas a drawing move lets both players force a draw. Although computing these types is generally infeasible in the opening and middlegame phases, endgames with up to six pieces have been definitively solved by modern computers. We can therefore measure the values of all endgame moves.

Chess also admits a natural measure of object complexity. As in any dynamic game, assessing the value of a chess move requires DMs to examine the ensuing subgame. Because larger subgames are likely harder to evaluate than smaller ones, we propose to equate the inherent complexity of a move with the size of the subsequent subgame. As an empirical proxy for the size of a subgame, we rely on what chess players refer to as depth-to-mate (DTM). DTM is a theoretical metric of how fast the dominant player can force a checkmate when the losing player resists as long as possible. In our data, moves' DTM is strongly correlated with the length of actual play, suggesting that DTM is a useful proxy for the size of the ensuing subgame and thus the difficulty of evaluating the corresponding move.

Our data come from two sources. The first one is `lichess.org`, one of the three most popular internet chess servers. We have information on the universe of moves in all rated games on the platform from January 2013 through August 2020.[2] We augment these data with information from the Syzygy and Nalimov endgame tablebases (Nalimov et al. 2000; Man 2013). Both databases record the type of every available move in all non-trivial endgame positions with up to six pieces—just as in Zermelo's Theorem. The Nalimov database also contains information on moves' DTM, our proxy for object complexity. Our analysis focuses on the choices of nearly a quarter million experienced users, ranging from seasoned hobbyists to the best players in the world. In total, we observe about 227 million choices from sets with approximately 4.6 billion alternatives.

The first part of our empirical investigation examines how objects' inherent complexity affects choice behavior. Comparing moves of the same type across otherwise similar choice sets, we find that winning moves become less likely to be chosen as their DTM increases. For losing moves, the reverse holds. Both findings are predicted by the theory.

To gauge the economic significance of object complexity, consider a two-node increase in the DTM of a winning move. According to our estimates, such an increase corresponds to a 1.0–2.3 p.p. decrease in the choice frequency of the corresponding move, which translates to a 10%–24% reduction relative to the mean choice frequency of winning moves.

[2]Rated games are consequential in the sense that their outcomes directly affect users' strength ratings and rankings on the site. Anecdotal evidence suggests players care intensely about their ratings.

We further document that, within the same the choice set, winning moves with a low DTM are more often selected than their counterparts with higher DTM. This pattern holds for the vast majority of players in the data. We prove that the observed pattern is consistent with satisficing and inconsistent with the predictions of maximization-based models in which the DM conducts noisy evaluations of all alternatives before choosing the one with the maximal estimated value. Our theory thus yields a new empirical test that leverages object complexity to distinguish satisficing from noisy maximization.

In the second part of the empirical analysis, we examine how skill and deliberation time moderate the adverse effect of complexity. Given that our data contain approximately 3.4 million moves by chess masters, of which nearly 350,000 were made by grandmasters, we can ask whether highly skilled players are less likely to make mistakes than less skilled ones, and how the difference between both kinds of players depends on complexity. By mistake we mean choosing a drawing or losing move in the presence of at least one winning alternative. In the data, titled players are significantly less likely to make mistakes than untitled ones; but only when the minimal DTM among winning moves is relatively high, i.e., when optimal moves are hard to evaluate. We also find that greater time pressure is associated with more mistakes, especially when the available moves are inherently complex. We further establish that when players are afforded more time to make decisions, titled players' frequency of suboptimal choices declines by more than that of untitled ones. Skill and deliberation time thus complement each other in moderating the impact of object complexity.

In the last part of the analysis, we examine set- rather than object-level factors that complicate decision-making. We show that the *mix* of suboptimal alternatives is another key determinant of suboptimal choice. Consider, for example, replacing a single losing move in the choice set with a drawing alternative. Based on our estimates, such a change increases the frequency of mistakes on average by about .84 p.p.—or 14% relative to the mean frequency of mistakes. Contrary to psychological notions of choice overload (Iyengar and Lepper 2000), we find that what matters for the quality of decision-making in this setting is the composition of the choice set, not its size. Once we carefully control for the fraction of winning, drawing, and losing moves in the set, the number of available alternatives and the frequency of mistakes are practically uncorrelated.

The work in this paper speaks directly to the theoretical literature on how complexity considerations affect outcomes in single- and multi-person environments. This research usually conceives of complexity as affecting behavior through constraints on agents' computational abilities and memory (e.g., Neyman 1985; Rubinstein 1986; Abreu and Rubinstein 1988; Kalai and Stanford 1988; Salant 2011; Wilson 2014; Jakobsen 2020). A high-level takeaway is that computational constraints can greatly affect both individual and strategic outcomes.

Our contribution relative to extant theoretical work is twofold. First, we develop a tractable model of choice that highlights objects' inherent complexity as a source of errors in decision-making. Second, we provide evidence on the empirical relevance of complexity considerations in general, and object complexity in particular.

In addition, our work complements a growing experimental literature on complexity in decision-making (see, e.g., Huck and Weizsäcker 1999; Gabaix et al. 2006; Abeler and Jäger 2015; Bossaerts and Murawski 2017; Enke and Graeber 2021). Rubinstein (2007, 2016), for instance, shows that decisions that require explicit cognitive reasoning take far longer to complete than those that do not. Caplin, Dean, and Martin (2011) provide evidence that individuals rely on satisficing in choice environments in which evaluating each option takes time and effort. Oprea (2020) develops a revealed-preference methodology to measure the cost of complexity. He finds subjects are willing to pay significant amounts in order to avoid tasks that are inherently complex. Overall, laboratory experiments support the idea that complexity can affect decision-making.

Outside of the laboratory, tests of fundamental decision-theoretic concepts remain rare.[3] As Chiappori, Levitt, and Groseclose (2002) note, nonexperimental settings may be intractable, with choice sets that need not be known in their entirety, or even be specified ex ante. Moreover, theoretical predictions may hinge on subtle properties of utility functions, intricacies of payoff structures, and individuals' beliefs—all of which are typically unobserved by the econometrician. As a result, we know little about how complexity affects decision-making in real-world environments.[4]

Chess endgames provide an almost ideal empirical setting to study this question. In addition to yielding observable variation in complexity and admitting an ordinal measure of alternatives' value, chess possesses at least four additional attractive features. First, the rules of the game are known to players and there is virtually no uncertainty about primitives such as choice sets. Yet, evaluating different moves often strains the bounds of human cognition. Second, data on chess games are abundant. We therefore have enough statistical power to test even subtle theoretical predictions. Third, chess games are played under different time controls, and there are large differences in players' skill. This allows us to analyze how complexity, skill, and time interact. Fourth, we study experienced players in a familiar environment, thus

[3]A related literature asks whether some of the basic tenets of game theory are consistent with observed behavior in different real-world environments. Walker and Wooders (2001), Chiappori, Levitt, and Groseclose (2002), Palacios-Huerta (2003), and Hsu, Huang, and Tang (2007) all study minimax play in professional sports, while Spenkuch, Montagnes, and Magleby (2018) examine backward induction in sequential voting. On the whole, the evidence from these settings corroborates theory more closely than one might have guessed based on an abundance of negative findings from the laboratory (see, e.g., Camerer 2003 for a review).

[4]Levitt and List (2007) argue that even if individuals consistently make mistakes in the laboratory, competitive forces and experience may limit such behavior in the real world.

minimizing the risk that our findings are due to an unfamiliar setting or driven by learning.[5]

Finally, scholars have long been intrigued by chess as a game that is theoretically trivial but practically intractable. Chess is a finite, two-player, zero-sum game with perfect information (Neumann 1928).[6] It is well known that, in any board configuration, either White has a winning strategy, Black possesses such a strategy, or both sides can guarantee themselves a draw (Zermelo 1913; Neumann and Morgenstern 1944). In the words of Osborne and Rubinstein (1994, p. 6):

> The existence of such strategies suggests that chess is uninteresting because it has only one possible outcome. Nevertheless, chess remains a very popular and interesting game. [...] Even if White, for example, is shown one day to have a winning strategy, it may not be possible for a human being to implement that strategy. Thus while the abstract model of chess allows us to deduce a significant fact about the game, at the same time it omits the most important determinant of the outcome of an actual play of chess: the players' "abilities." Modeling asymmetries in abilities and in perceptions of a situation by different players is a fascinating challenge for future research, which models of "bounded rationality" have begun to tackle.

To this we only add that it is also important to *empirically* explore deviations from optimal play; and our study can be understood as taking a step in this direction.

2. Theory

Our analysis begins by developing a general model of satisficing with evaluation errors. After specializing the model to the case of chess and deriving its empirical predictions, we explain how to leverage object complexity to empirically distinguish between satisficing and standard forms of utility maximization.

2.1. *General Setup*

A decision maker (DM) chooses a single object from every choice set $A \subseteq X$, where X is the grand set of all feasible alternatives. Each object $x \in X$ is characterized by a pair (v_x, σ_x), where v_x denotes the value of x to the DM and σ_x corresponds to its inherent complexity. The latter can be thought of as the length of an efficient representation of the object.

When the DM encounters an alternative x, she obtains a noisy estimate of value:

$$u_x = v_x + \epsilon_x, \tag{1}$$

[5]For conflicting evidence as to whether experience and skill in one strategic environment transfer to another one, see Palacios-Huerta and Volij (2008, 2009), Wooders (2010), and Levitt, List, and Reiley (2010); Levitt, List, and Sadoff (2011).

[6]There is a question as to whether chess should, in fact, be considered a finite game (see, e.g., the discussion in Osborne and Rubinstein 1994, p. 100). At least since 2014, the FIDE rule book specifies that a game is *automatically* drawn if the same position has occurred five times (cf. Article 9.6.1). Under official rules, chess is thus finite.

where the evaluation error ϵ_x is distributed normally with mean zero and standard deviation σ_x. Thus, the DM needs to contend with noise when evaluating a particular alternative, and the amount of noise increases in the object's inherent complexity. The DM's assessments are only correct in expectation.[7]

In choosing from choice set A, the DM wishes to pick a high-value object while economizing on effort. To this end, she follows a satisficing procedure whereby she selects the first "satisfactory" alternative she encounters. Formally, the DM has in mind an aspiration level T. She randomly draws an object from A and evaluates it. The object is chosen if its estimated value exceeds T. Otherwise, the DM discards the object as non-satisfactory and draws the next one (without replacement). She continues in this fashion until she finds a satisfactory object or until she exhausts all alternatives in A without making a choice. In the latter case, the DM chooses the last object that she examined.[8]

The outcome of this choice procedure can be summarized by a random choice function C that maps every choice set A and every $x \in A$ to the probability $C(x, A)$ of selecting x from the set A. Choice behavior is stochastic because object evaluations are noisy and because the evaluation order is random.[9]

2.2. *Application to Chess*

In chess, the set X is the set of all possible legal moves of all pieces, and a choice set corresponds to all legal moves in some board configuration. Players' choice sets can thus be directly mapped to board configurations.

By Zermelo's Theorem, starting from any given configuration, either White can force a win, Black can force a win, or both sides can guarantee themselves a draw. It is therefore possible to associate every move in any board configuration with the ultimate outcome of the game under subsequent optimal play. A move that allows the DM to force a win yields positive payoff W, whereas a move that enables her opponent to do so produces a negative payoff

[7] Two key factors that may mitigate the adverse effect of complexity on value assessments are skill and deliberation time. We postpone the discussion of these moderators to Section 6, where we analyze the relevant empirical regularities, and enrich the model to accommodate them.

[8] If the DM recalls the entire history of u's then she may choose the u-maximal object. This form of recall complicates the analysis to follow without adding substantive insights. Intuitively, this is because in the large sets we consider and assuming the threshold is not too high, exhausting all alternatives without exceeding T is a low probability event with a second-order effect on the theory's empirical predictions.

[9] The evaluation order does not favor any of the alternatives. Conditional on not stopping and not having been evaluated yet, every alternative x is equally likely to be examined next, and its conditional choice probability is determined by v_x and σ_x. If the evaluation order is allowed to favor some alternatives over others, then *any* random choice function C can be rationalized by a choice procedure that *(i)* draws the first alternative to be evaluated with its choice probability, and *(ii)* always chooses the alternative that is evaluated.

$L = -W$. Moves that lead to draws generate a payoff of $D = 0$.[10] Hence, $v_x \in \{W, D, L\}$ for every move x. As a matter of terminology, we say move x in some choice set is of type W, or simply a W-move, if $v_x = W$, with analogous designations for moves that yield payoffs of D and L.

Our analysis focuses on choice sets that include at least one W- and at least one D- or L-move. In these sets, it is clear what it means to choose suboptimally. Moreover, the value of the maximal alternative remains unchanged as moves are added to the choice set. Both of these features are helpful in the empirical analysis.

We make two further assumptions. The first one guarantees that winning moves, when accurately estimated, are satisfactory.

Assumption 1: *The threshold T is between D and W.*

The second assumption is only relevant for Predictions 1 and 4. It states that the inherent complexity of a D-move is not "too small" relative to that of L-moves. In Section 3, we argue that this assumption is plausible in the context of chess.

Assumption 2: *In any choice set, the inherent complexity of any D-move is at least half of the inherent complexity of any L-move.*

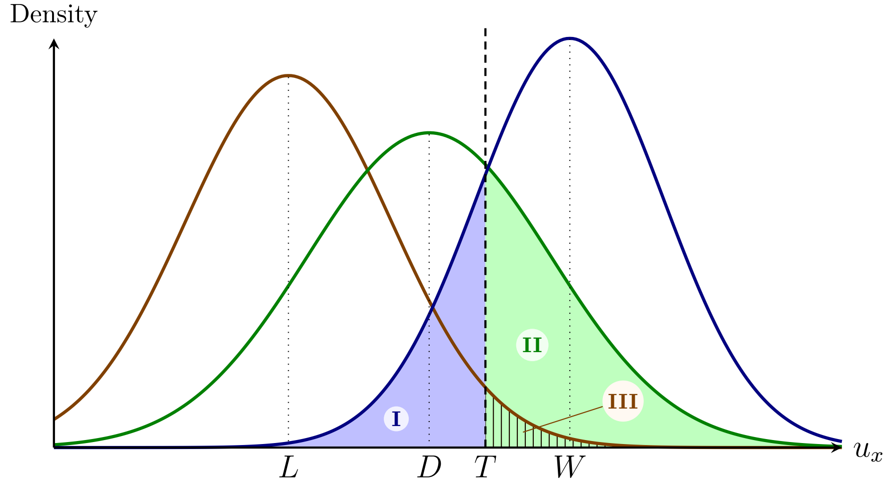
Under these assumptions, the model predicts that the choice probabilities of individual moves increase in their values.

Prediction 1: *The choice probability of every W-move in a choice set exceeds the choice probability of every D-move in the same set, which in turn is larger than the choice probability of every L-move in the set.*

Figure 1 provides a graphical intuition for this and subsequent predictions. Area I in this figure corresponds to the probability that the DM discards a W-move, conditional on encountering one. Areas II and III correspond to the probability of accepting a D- or an L-move. Given that $W > T > D$ and in light of the symmetric distribution of evaluation errors, areas I and II must each be smaller than one half. This observation, combined with random sampling of alternatives, means W-moves are chosen more frequently than D-moves. Figure 1 also makes clear why the inherent complexity of D-moves cannot be "too small" relative to moves of type L for Prediction 1 to hold. If it were, then area III may be larger than area II, which

[10] The assumed payoff structure reflects the fact that chess is usually considered a zero-sum symmetric game. This payoff structure is not critical for our analysis. For the predictions below, it suffices to assume that $W > D > L$ and modify Assumption 2 appropriately.

Figure 1: Stochastic Evaluation of Moves



Notes: Figure illustrates the evaluation process in our model, as applied to chess.

would imply that the choice probability of the corresponding L-move is larger than that of the corresponding D-move.

2.3. *Object Complexity*

The model makes two predictions regarding the connection between object complexity and choice behavior. The first prediction pertains to changes in the complexity of a single move, holding all else equal.

PREDICTION 2: *As the complexity of a W-move increases, its choice probability decreases and that of all other available moves increases. As the complexity of an L- or D-move increases, its choice probability increases and all other choice probabilities decrease.*

To build intuition, recall that higher object complexity corresponds to noisier evaluations of the respective moves and thus to "wider" distributions in Figure 1. Hence, an increase in the complexity of a W-move leads to more probability mass in area I, which implies a lower choice probability. As for the remaining moves' choice probabilities, they move in the opposite direction because the evaluation order is random, and in any given order, the choice probabilities of all subsequent moves must offset the change in the choice probability of the current move. For D- or L-moves, an increase in object complexity leads to more probability mass in area II or III, and thus to opposite comparative statics.

An immediate implication of Prediction 2 relates to suboptimal choice. We say that a DM chooses suboptimally, or makes a mistake, when she picks a D- or an L-move in the presence of at least one W-alternative. By Prediction 2, as the complexity of an optimal move increases, some choice probability shifts from that move to suboptimal ones, whereas an increase in the complexity of a suboptimal move leads to a shift in choice probability to optimal alternatives. Therefore:

COROLLARY 1: *Mistakes become more likely as the object complexity of any move in the choice set increases.*

The next prediction compares choice probabilities within the same set.

PREDICTION 3: *Fix a choice set A and two moves $x, y \in A$ with $\sigma_x < \sigma_y$. If both are W-moves then the choice probability of x is larger than that of y. However, if both are D- or L-moves then the choice probability of x is smaller than that of y.*

Intuitively, a more complex W-move has more probability mass below the threshold than a simpler one, which implies that it is chosen with smaller probability. For two D- or two L-moves, larger complexity implies more mass above the threshold, and hence the opposite ranking of choice probabilities.

2.4. *Satisficing vs. Maximization*

Our satisficing-with-evaluation errors model combines two related but conceptually distinct ideas. First, we stipulate a noisy evaluation process that resembles the familiar random-utility framework in discrete-choice models (Luce 1959; Marschak 1960; McFadden 1974). An important difference between the evaluation process in our model and the standard discrete-choice setup is that ϵ_x in eq. (1) corresponds to an error in perception. Moreover, the variance of the error term depends directly on the object's inherent complexity. This feature is essential for generating predictions on how object complexity affects choice.

The second idea is to incorporate noisy evaluations into a satisficing choice procedure. Drawing directly on Simon's (1955; 1972) discussion of bounded rationality and satisficing in chess, we model the DM as searching for and accepting the first "good enough" move. According to Simon (1955, p. 100), "in actual human decision-making, alternatives are often examined sequentially. [...] If a chess player finds a forced mate for his opponent, he generally adopts this alternative without worrying whether another alternative also leads to a forced mate."

Would the model yield similar predictions if it featured a standard maximization procedure rather than satisficing? By standard maximization, we mean any choice procedure in which

- the DM considers all moves in the choice set, obtains estimates of their values, and forms, for each move, a posterior belief about its value;

- the ranking of moves' expected values according to the posterior coincides with the ranking of the estimates; and
- the DM chooses the move with the highest posterior expected value.

Such a procedure encompasses a large class of models. Perhaps most importantly, it includes the standard "$max_x u_x$"–approach in the discrete-choice literature. It also includes some models of noisy cognition, in which agents choose the alternative with the highest Bayesian posterior mean (see, e.g., Woodford 2020 for a review).

Standard maximization and satisficing yield opposing predictions regarding the effect of object complexity on choice probabilities. Specifically, under maximization, the reverse of Prediction 3 holds.

PROPOSITION 1: *Assume standard maximization as choice procedure, and consider any choice set A with two W-moves $x, y \in A$ such that $\sigma_x < \sigma_y$. The choice probability of x is smaller than that of y.*

The intuition for Proposition 1 is as follows. Although the distribution of the difference $u_y - u_x$ is symmetric with mean zero, the distribution of u_y has "thicker" tails and, most importantly, a thicker right tail than that of u_x. The thicker right tail implies a positive probability of u_y being greater than u_x conditional on both of them exceeding "large" thresholds. Under maximization, the threshold they need to exceed is the maximum among all other moves and so, in expectation, it is large.

Proposition 1 and Prediction 3 provide the foundation for a new empirical test that distinguishes satisificing from utility maximization. This test is applicable in any environment in which there are at least two optimal alternatives that differ from one another in their inherent complexity. Satisficing predicts that the less complex optimal alternative will be chosen more often whereas maximization predicts the reverse.

3. Data

Taking the theory to the data requires information on actual choice behavior, as well as the available alternatives' values and object complexity. In this section, we describe how we assemble a data set on chess endgames that contains this information.

Information on choice behavior comes from `lichess.org`, one of the three most popular online chess platforms. Funded by donations, Lichess is ad free and allows anyone to play live chess games at no cost through a high-quality graphical user interface (see Figure 2 for a screenshot of a typical game). Although Lichess offers a choice between many different time limits, the majority of games that are actually hosted on the platform can be broadly classified as "speed chess," i.e., games in which players have significantly less time to complete

Figure 2: Screenshot of a Rated Game on Lichess

Notes: Figure shows a screenshot from a rated game between registered users on `lichess.org`. The green squares highlight the most recent move, i.e., ♙c4.

their moves than under classical time controls.[11] Lichess further distinguishes between casual and rated games. The latter determine player ratings and are therefore only available to registered users. In a nutshell, a player's strength rating increases (decreases) whenever she wins (loses) a rated game, and it increases (decreases) by more the stronger (weaker) her opponent was.[12] Anecdotal evidence from online messaging boards suggests that users often care intensely about their rating. Since high ratings tend to be a source of pride among chess players, Lichess has a strict policy against computer-assisted play. Enforcement of this policy relies on a variety of methods, including community reporting of suspected offenders and automatic detection algorithms.

We have data on the universe of rated games between human players from January 2013 through August 2020. The available information includes players' usernames, ratings and real-world titles, if any (i.e., candidate master, national master, international master, grandmaster, etc.), the date and start time of the game, its outcome, and the precise sequence of moves.[13] We can therefore reconstruct all choice sets that a player faced as well as the moves she chose.

We augment these data with information on moves' types and, for W- and L-moves, their depth-to-mate (DTM). In principle, this information is computable for any legal move in any

[11]Some of our analyses restrict attention to games with classical controls, in which time pressure is less of an issue. As a robustness check, we also estimate alternative specifications that control for how much time a player has left on the clock. The results are qualitatively equivalent to those below (cf. Appendix C).

[12]In calculating ratings, Lichess implements the Glicko-2 algorithm, which extends the better-known ELO system by incorporating a measure of uncertainty in approximately Bayesian fashion (Glickman 1999, 2000).

[13]Lichess verifies that users do, indeed, hold the real-life titles they claim.

chess game, but doing so is computationally infeasible in the opening and middlegame phases. Our analysis thus focuses on endgame configurations with six or fewer pieces, which have been definitively solved by computer algorithms.[14]

These algorithms begin by constructing an exhaustive list of all possible (up to symmetry) legal board configurations with three chess pieces.[15] Every configuration is examined, and the ones in which the player to move is in checkmate are stored as "mated in 0." Next, all configurations with the other side to move are evaluated. If one of them can reach a configuration that has previously been determined to be "mated in 0" by executing a legal move, then it is stored as "mate in 1." To find the set of configurations that are "mated in 2," the algorithm looks for configurations from which *all* possible legal moves lead to "mate in 1" configurations; and to determine configurations that are "mate in 3," it subsequently checks for configurations from which it is possible to directly reach a configuration that is known to be "mated in 2." Proceeding recursively, a configuration is classified as "mated in l" if and only if it is impossible to avoid moving to configurations that are "mate in w," with $w = l - 1$ for at least one available legal move and $w \leq l - 1$ for all others. By contrast, a configuration is marked as "mate in w" if it is possible to move to another one that is "mated in $w - 1$." This procedure continues until no further progress at classifying configurations is made, at which point all remaining configurations with three chess pieces are designated as "drawn." Essentially the same algorithm is next applied to board configurations with four pieces, then five, and then six.

The end result is a so-called tablebase in which *board configurations* are classified as either "drawn," "mated in l," or "mate in w." A particular *move* is said to be of type W with DTM d if it results in a new board configuration that, with the other player to move, is known to be "mated in $d - 1$." Thus, the minimal DTM among all available W-moves from any configuration that is "mate in w" is, by construction, equal to w. Similarly, a move is said to be an L-move with DTM d if it leads to a configuration that is "mate in $d - 1$." The maximal DTM among all L-moves from any configuration that is "mated in l" equals l. Moves that result in "drawn" configurations are classified as type D.[16]

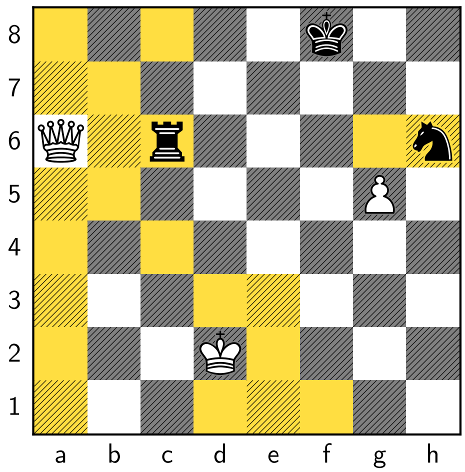
For a concrete example of the content of such a tablebase, consider Figure 3. The left panel depicts the board configuration that is being examined, with the data for each available legal move shown on the right. The assessment of a move consists of two components: its

[14] Although board positions with seven pieces have, in principle, been solved as well, we do not rely on them in the analysis, because doing so would be impractical. Specifically, the open-source Syzygy tablebases do not contain information on DTM. The commercially available seven-piece Lomonosov tablebases do contain information on DTM but require about 140TB of storage. They are thus too large to be usable in most computing environments.

[15] It is not necessary to consider configurations with two lone kings, as these are automatically drawn.

[16] For additional information on algorithmic analysis of chess endgames, see, e.g., Thompson (1986).

Figure 3: Example of an Endgame Table



Move	Evaluation
♕c6	W in 19
♕a8	W in 21
♕a3	W in 47
♕a7	D
♕a5	D
♕a4	D
♕a2	D
♕a1	D
♕b5	D
♕d3	D
♕e2	D
♕f1	D

Move	Evaluation
♕b7	D
♙h6	L in 34
♕c4	L in 34
♕c8	L in 34
♔e2	L in 32
♔e3	L in 32
♔d3	L in 32
♕b6	L in 32
♔e1	L in 30
♔d1	L in 28
♙g6	L in 28

Notes: Figure provides an example of the information in endgame tablebases. The left panel shows the board configuration that is to be evaluated, assuming it is White's turn. Yellow-colored squares help visualize the set of available moves. The right panel shows the computer evaluation of each legal move, drawing on the Nalimov endgame tables. W, D, and L denote win-, draw-, and loss-moves from the perspective of the current player.

type (i.e., W, D, or L), and, for a W- or L-move, its DTM. In this particular example, ♕c6 corresponds to "W in 19," which means that, if White moves the queen to c6, then White can force checkmate in nineteen moves regardless of Black's response.

Endgame tablebases do not contain a measure of subgame depth for D-moves. To the best of our knowledge, there is no general approach to even identify D-moves by means other than elimination, which is how computer algorithms proceed. A classification-by-elimination approach, however, does not lend itself to a natural measure of complexity for D-moves—other than suggesting it may be large relative to that of other alternatives in the choice set. While one could, in principle, calculate the depth of the subgame associated with a particular D-move by iteratively developing the game tree, doing so at scale is computationally infeasible. We therefore refrain from quantifying object complexity for D-alternatives, and restrict our empirical tests of Predictions 2 and 3 to W- and L-moves.[17]

Given the results of extant computer analyses, the underlying value, or type, of essentially any endgame move is known with certainty, as is the DTM of W- and L-moves.[18] We retrieve the relevant information for every legal move in every endgame configuration in the Lichess

[17]Prediction 1 relies on Assumption 2, which intuitively says that identifying a D-move with certainty is at least half as difficult as identifying L-moves. Per the discussion above, this assumption may be reasonable. It is also reasonable to interpret our empirical test of Prediction 1 as a joint test of the prediction and Assumption 2.

[18]The only exceptions are positions with castling rights and configurations in which a lone king faces five other pieces. The former are extremely rare in endgames ($< .01\%$ of available legal moves in our data), while the latter are uninteresting (because 98.8% of available moves are of type W). Our empirical work excludes all board configurations for which information on DTM is not available.

data by running several billion queries against the Syzygy and Nalimov tablebases (Nalimov et al. 2000; Man 2013).[19]

Our final sample contains nearly 227 million decision problems with a total of over 4.6 billion legal moves. There are four distinct sources of selection into this sample. First, we restrict attention to choice sets that contain at least one W- and at least one D- or L-move. We impose this restriction for expositional convenience. When a board configuration admits at least one W-move, any choice of a D- or L-alternative is unambiguously a mistake. In addition, we can distinguish "small" mistakes from "large" ones. The former correspond to picking a D-move in the presence of one or more of type W, while the latter consist of choosing an L-move instead. A disadvantage of this restriction is that the choice sets in our sample contain an above average share of W-moves.

The second and related source of selection pertains to how often different individuals reach a winning position, i.e., an endgame position with at least one W-move. The strongest players, for instance, may often mate their opponents before reaching the endgame stage. Similarly, very weak players may rarely be in a position to win endgames and might thus also be underrepresented in our sample. In what follows, we address this issue in two ways. First, whenever appropriate, we control for player fixed effects. Second, we reweight observations so that all players receive equal weight in our analysis. The results below should hence be interpreted as referring to a typical decision by the average player in our sample.[20]

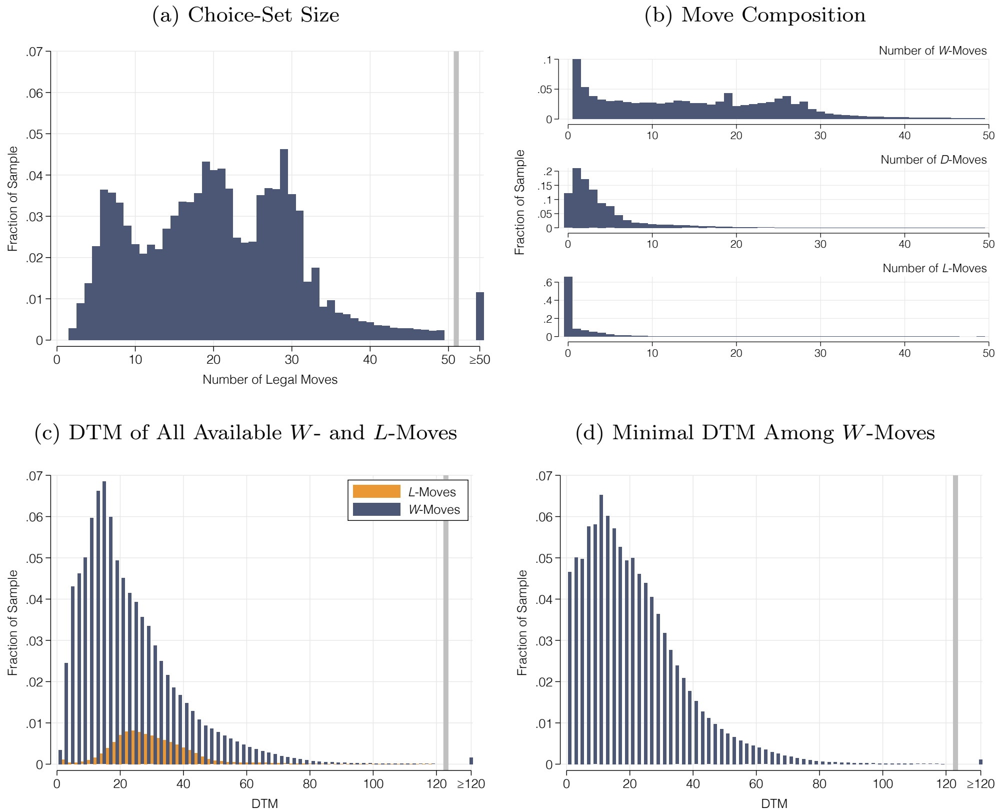
Third, to minimize the risk that our findings are due to an unfamiliar setting or a lack of experience with similar decision problems, we exclude, for every player, the first one thousand endgame moves from winning positions. We are thus left with highly experienced DMs who are very familiar with the task at hand.

Finally, users on Lichess are not a random subset of all experienced chess players. In the appendix, we address this potential source of concern by replicating our main results in an independent data set covering a large number of chess games in international tournaments. These data come from the online publication *The Week in Chess* (TWIC), which covers "all the latest news and games from international chess." The main disadvantage of this alternative data set is that there is less variation in the skill of players, and that it is more than two orders of magnitude smaller. These limitations notwithstanding, the TWIC data yield qualitatively similar conclusions (cf. Appendix D).

[19] As a technical side note, the Syzygy tablebases do not contain information on DTM. In contrast to the Nalimov tables, they do, however, take into account the fifty-move stalemate rule. In rare instances, the fifty-move rule matters for correctly determining whether one player can unilaterally invoke a draw. We therefore rely on win-draw-loss information from the Syzygy database, while information on DTM comes from Nalimov's database.

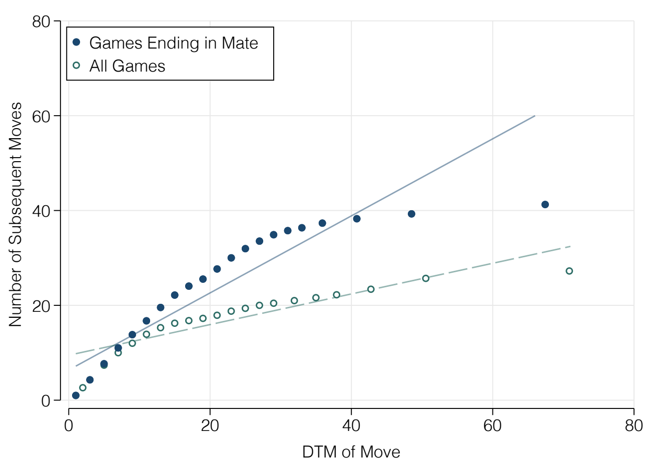
[20] For robustness checks in which every individual observation receives the same weight, see Appendix C.

Figure 4: Distributions of Key Variables



Notes: Panel (a) presents the distribution of the number of available moves, with panel (b) distinguishing between W-, D-, and L-moves. Panel (c) shows the distribution of DTM among W- and L-moves, and panel (d) depicts the distribution of the smallest DTM among all W-moves in the choice set.

Figure 5: Greater Depth to Mate is Associated with Longer Realized Game Paths



Notes: Figure shows binned scatter plots of the relationship between chosen moves' DTM (x-axis) and the total number of subsequent moves in the same game (y-axis). Hollow circles correspond to within-bin averages for all games, while solid circles restrict attention to games in which one of the players is eventually mated. The solid and dashed lines show the lines of best fit in the respective underlying move-level data.

4. Descriptive Statistics

We next describe some of the basic features of our data, starting with the distributions of key variables. The upper two panels in Figure 4 display histograms for the total number of available legal moves in the DM's choice set (left) and for the number of W-, D-, and L-moves (right). Choice sets contain, on average, about 20.6 moves, of which 15.5 are W-moves. There is significant variation in choice-set size ranging from less than a handful of moves to more than thirty, or even fifty alternatives. Most, but by no means all, of this variation is due to differences in the availability of W-moves. The data therefore allow us to study choices from small, medium, and large sets.

The lower two panels of Figure 4 demonstrate that there is significant variation in our measure of object complexity. The lower-left panel depicts a histogram of the DTM of all available W- and L-moves. The modal DTM is 15 among the former and 24 among the latter. Given the size of the sample, there are millions of legal moves with DTM of five or less, and millions of moves with DTM of more than fifty.

The lower-right panel plots the distribution of the minimal DTM among W-moves at the choice-set level. From a theoretical perspective, minimal DTM corresponds to the minimal amount of noise with which DMs need to contend in order to identify the value of at least one W-move. Empirically, the correlation between minimal DTM and other low-dimensional proxies for the available moves' complexity, such as the average or median DTM in the set, exceeds 0.8. The histogram in this panel indicates that the data include choice sets in which evaluating some of the moves is relatively easy, others where accurately classifying optimal moves likely exceeds the bounds of human cognition, and a great range of intermediate cases.

Figure 5 presents a binscatter plot of the relationship between moves' theoretical DTM and the total number of subsequent moves in the same game. The former may differ from the latter for a variety of reasons. For example, if the dominant player succeeds in mating her opponent but does not take the shortest path to victory, then the total number of subsequent moves may exceed the initial move's DTM. If, however, the losing player resigns or does not hold out as long as possible, then there will be fewer subsequent moves than implied by DTM. Importantly for our purposes, we observe a monotonically increasing relationship between moves' DTM and the length of the realized path of play. In other words, higher-DTM moves are, indeed, associated with longer subgames.

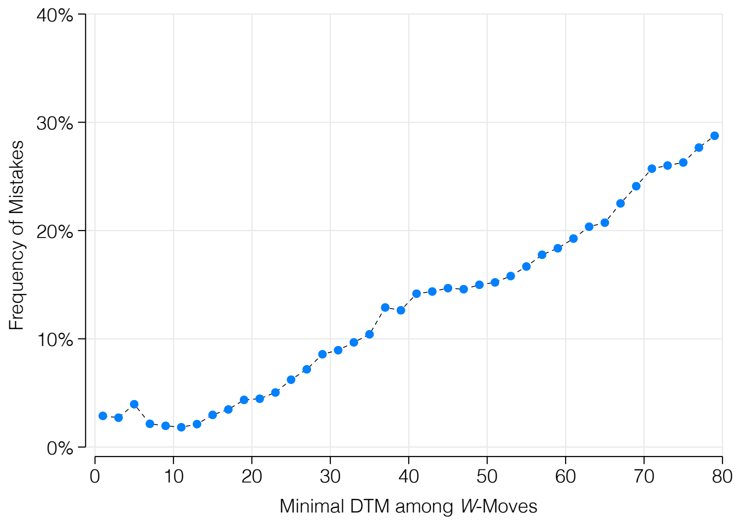
Some of our analysis focuses on suboptimal choice and mistakes. By mistake, we mean choosing a D- or an L-move in the presence of W-alternatives. The summary statistics in Table 1 show that about 6% of observed moves are mistakes. Given that, on average, W-moves account for about 75% of all available moves in a choice set, mistakes occur at nearly one quarter the rate one would expect if DMs were choosing at random. There is thus evidence

Table 1: Summary Statistics

Variable	Mean	SD	Percentile 25%	50%	75%	95%	N
A. Move Characteristics							
Type:							
W-Move	0.69	0.46					4,617,441,573
D-Move	0.23	0.42					4,617,441,573
L-Move	0.08	0.27					4,617,441,573
DTM:							
W-Moves	25.89	17.88	13	23	33	59	3,457,878,398
L-Moves	30.35	13.46	22	28	36	50	296,522,573
B. Choice-Set Characteristics							
Set Composition:							
Total Number of Legal Moves	20.58	10.35	13	20	28	38	226,955,095
Number of W-Moves	15.48	11.09	6	15	24	34	226,955,095
Number of D-Moves	3.81	4.30	1	2	5	13	226,955,095
Number of L-Moves	1.29	2.73	0	0	2	7	226,955,095
DTM of W-Moves:							
Shortest	20.91	16.18	9	17	29	51	226,955,095
Median	25.81	17.34	13	22	33	59	226,955,095
Longest	31.62	20.05	17	27	39	69	226,955,095
DTM of L-Moves:							
Shortest	26.95	11.39	20	26	34	44	77,045,730
Median	30.19	12.50	22	28	36	48	77,045,730
Longest	34.27	16.31	24	32	40	60	77,045,730
C. Outcomes							
Mistakes:							
Any Type of Error	0.06	0.24					226,955,095
Small Mistake	0.05	0.23					226,955,095
Large Mistake	0.01	0.08					226,955,095
Result of Game:							
If Current Move is Mistake:							
Win Game	0.31	0.46					13,052,773
Draw	0.49	0.50					13,052,773
Lose Game	0.19	0.40					13,052,773
If Choose W-Move:							
Win Game	0.74	0.44					213,902,322
Draw	0.20	0.40					213,902,322
Lose Game	0.05	0.22					213,902,322
D. Timing							
Time Left on Clock (in sec.)	72.35	192.87	8	22	69	296	212,295,223
Deliberation Time (in sec.)	1.66	3.07	0	1	2	5	212,249,738
E. Player Characteristics							
Total Number of Endgame Moves	2,584	2,469	1,297	1,793	2,877	6,696	237,232
Average Rating	1,733	281	1,533	1,716	1,917	2,222	237,232
Real-World Title	0.01	0.10					237,232

Notes: Table displays summary statistics for selected variables in the data. Each observation in panel A corresponds to a legal move, and observations in panels B–D correspond to decision problems. Panel E contains player-level information. Observations are reweighted so that all players receive equal weight. The number of observations related to the timing of moves is smaller because the raw data do not include this information for games that were played prior to April 2017.

Figure 6: Greater Object Complexity is Associated with More Suboptimal Choice



Notes: Figure shows the relationship between the frequency of mistakes (y-axis) and the minimal DTM among available W-moves (x-axis).

of significant skill among the players in our data. Yet, the raw data also imply that mistakes occur with a certain regularity. They are not rare events. Moreover, mistakes tend to be consequential. A player whose current move is a mistake is about 43 p.p.—or roughly 58%—less likely to ultimately win the game than one who chooses a W-move. At the same time, the probability of a loss more than doubles.[21]

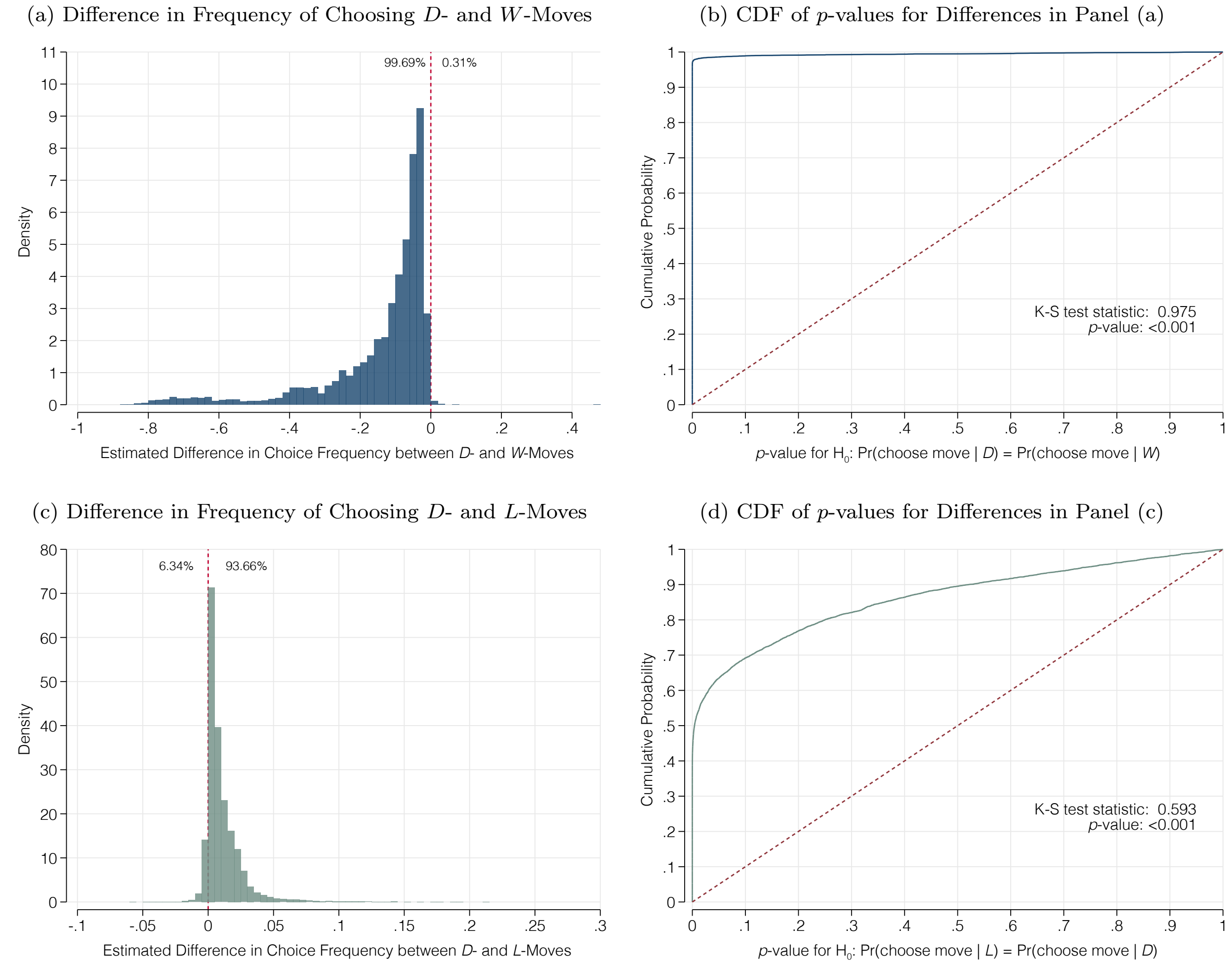
We also observe that DMs are more likely to make mistakes when the available moves are inherently more complex. Figure 6 illustrates this point by plotting the frequency of mistakes against the minimal DTM among winning moves. As explained above, minimal DTM is highly correlated with other low-dimensional measures of objects' complexity, such as the mean or median DTM. Regardless of which summary measure we ultimately use, suboptimal choice is strongly related to the available moves' inherent complexity.

To study choices at a more granular level, we partition the data into classes of choice sets so that two sets are in the same class if and only if they have the same number of W-, D-, and L-moves. We observe 9,701 such classes with more than a hundred choices. For each class, we calculate the empirical frequency with which a legal move of a particular type is chosen, and plot the observed differences in Figure 7.

The top-left panel shows a histogram of differences in choice frequency between D- and W-moves—one estimate for each class—and the bottom-left panel does so for the difference between D- and L-alternatives. To the right of each histogram, we show the CDF of the

[21]Mistakes do not always result in forgone wins because the player's opponent may subsequently also choose a suboptimal move. Similarly, due to the potential for mistakes in future moves, choosing a W-move in a particular instance does not automatically mean that the player will ultimately win the game.

Figure 7: Comparing Choice Frequencies of W-, D-, and L-Moves



Notes: Figure compares choice frequencies of individual W-, D-, and L-moves. Panel (a) plots the distribution of estimated mean differences in choice frequencies of D- and W-moves in choice sets with identical composition of moves, conditional on observing at least 100 choices from sets with that makeup. Negative estimates imply that, in the respective classes of choice sets, D-moves are, on average, less frequently chosen than W-alternatives. Panel (b) depicts the empirical CDF of the two-sided p-values associated with the estimated differences in panel (a). It also shows results from a Kolmogorov-Smirnov test against the null hypothesis of a uniform distribution of p-values. Panel (c) mirrors panel (a) but focuses on the difference in choice frequencies between D- and L-moves. Positive estimates imply that D-moves are, on average, more frequently chosen than L-alternatives. Panel (d) presents the empirical CDF of p-values for the null hypothesis that, in choice sets with a particular composition, L-moves are as frequently chosen as D-moves. All p-values account for two-way clustering by player and game.

associated p-values. Under the null hypothesis that choice probabilities do not depend on moves' types, the p-values of the estimated differences should be approximately uniformly distributed over the unit interval. This is not the case. In the vast majority of classes of choice sets, different types of moves are chosen with frequencies that are statistically distinguishable at conventional significance levels. More importantly, the signs of the observed differences are generally as postulated by the theory. That is, W-moves are more likely to be chosen than D-moves in over 99% of classes, and D-moves are more likely to be chosen than L-moves in over 93% of classes. Prediction 1 is thus borne out in the data.

5. Object Complexity and Choice Frequencies

We now proceed to test the two main predictions of the theory regarding object complexity and choice. Prediction 2 is concerned with how choice probabilities change as, all else equal, evaluating a particular legal move becomes more difficult. Because every board configuration is associated with a fixed choice set, the relevant empirical comparisons are necessarily across decision problems. By contrast, Prediction 3 compares the choice probabilities of different moves of the same type within the same choice set. According to Prediction 3, we should see that the W-move with the smallest DTM is chosen more frequently than the W-alternative with the second-smallest DTM value, which, in turn, is picked more often than that with the third-smallest value, and so on. For L-moves, the opposite pattern should emerge.

5.1. *Across-Choice-Set Comparisons*

To investigate Prediction 2, we focus on W- and L-moves and estimate the following econometric model:

$$Choose_a = \omega DTM_a \times W_a + \iota DTM_a \times L_a + \pi W_a + \phi_{A\setminus a} + \mu_p + \varepsilon_a. \tag{2}$$

Here, $Choose_a$ is an indicator for whether player p facing choice set A picked move a, DTM_a denotes that move's depth to mate, W_a and L_a are indicators for its type, and μ_p is a player fixed effect. We also include $\phi_{A\setminus a}$, a fixed effect for the other moves in the choice set. In constructing this fixed effect, we assume that, in line with the theory, moves can be reduced to their types and inherent complexity. Since we do not measure object complexity for D-moves, $\phi_{A\setminus a}$ conditions (only) on the vector of DTM values for W- and L-moves and the type composition of the choice set, i.e., the number of W-, D-, and L-alternatives. By including $\phi_{A\setminus a}$, we aim to approximate the thought experiment in which we vary the object complexity of one particular alternative, holding the values and object complexity of all other moves fixed.

Table 2 shows results from estimating variants of the regression model in eq. (2). In the first three columns, we study W- and L-moves from all board configurations. In the last three columns, we restrict attention to choice sets that do not contain any D-moves. The idea that our fixed effects appropriately control for the object complexity of all other moves is most plausible in the latter set of specifications. Reassuringly, the results in both sets of columns are qualitatively similar. Regardless of which sample we consider, we find that individual W-moves are significantly *less* likely to be chosen as object complexity increases. By contrast,

Table 2: Choices as a Function of DTM

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of Choosing Move					
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM × W-Move (÷ 100)	–1.140		–0.974	–0.521		–0.333
	(0.002)		(0.002)	(0.004)		(0.003)
DTM × L-Move (÷ 100)		0.034	0.007		0.066	0.054
		(0.001)	(0.001)		(0.002)	(0.002)
W-Move (÷ 100)			63.450			41.857
			(0.069)			(0.114)
Fixed Effects:						
Number of W- × D- × L-Moves × Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	9.397	0.785	8.495	10.881	1.183	8.078
R^2	0.372	0.209	0.368	0.513	0.209	0.494
N	3,457,878,398	296,522,573	3,754,400,971	398,856,135	111,905,262	510,761,397

Notes: Entries are coefficients and standard errors from estimating ω, ι, and π in variants of eq. (2) by ordinary least squares. All regressions control for player fixed effects as well as a fixed effect for the combination of the number of W-, D-, and L-moves and the vector of DTMs of all other W- and L-moves in the same choice set, as explained in the text. The unit of observation in each regression is an available legal W- or L-move, and observations are reweighted so that all decision problems for a particular player and all players receive equal total weight in the full sample of moves. The sample in cols. (1)–(3) includes all board configurations in our data, whereas cols. (4)–(6) restrict attention to configurations for which the associated choice sets do not contain D-moves. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are two-way clustered by player and game, and are shown in parentheses.

the choice probabilities of L-moves *increase* in their inherent complexity.

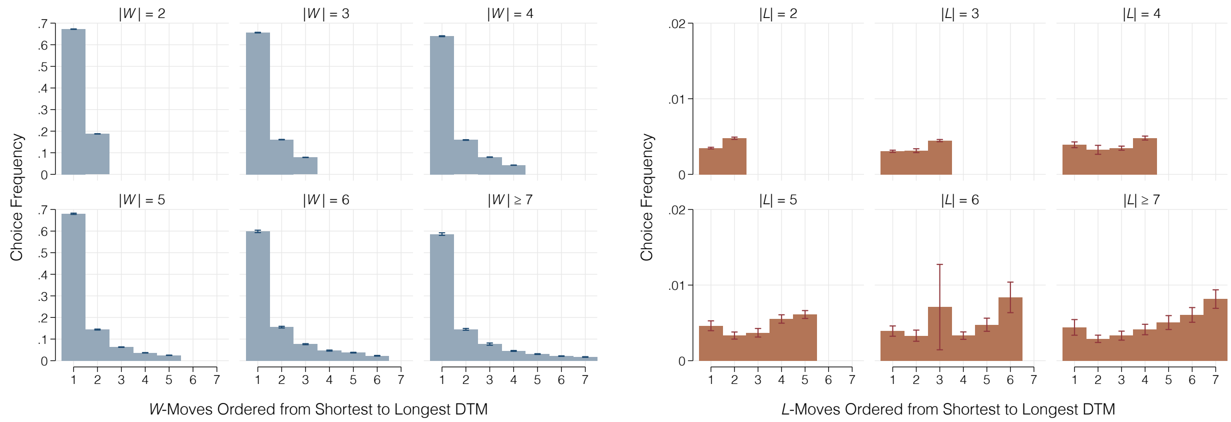
The estimates in Table 2 are not only directionally consistent with Prediction 2, but are economically large. According to the coefficients in cols. (1) and (4), increasing DTM by two nodes decreases the choice probability of the corresponding W-move by 1.0–2.3 p.p. This decrease translates to a 10%–24% reduction relative to the mean choice probability of W-moves. Our results therefore suggest that object complexity is an empirically meaningful determinant of choice.

5.2. *Within-Choice-Set Comparisons*

Our examination of Prediction 3 begins by looking at the raw data. We first partition choice sets into classes according to the number of available W-moves. In the left panel of Figure 8, we then plot, for each class, the empirical frequency with which DMs choose the W-move with the smallest DTM, the second-smallest DTM, and so on. We observe large and precisely estimated differences in choice frequencies. In line with Prediction 3, these frequencies decrease monotonically in moves' relative complexity, and all differences are statistically significant at conventional levels.

Figure 8: Choice Frequencies, by Relative Object Complexity

(a) W-Moves (b) L-Moves



Notes: Figure compares choice frequencies of W- and L-moves according to the ordering of their DTMs relative to other moves of the same type in the same choice set. Panel (a) does so for W-moves, and panel (b) focuses on L-moves. Each plot within either panel restricts attention to choice sets that contain a particular number of W- or L-moves with distinct DTM values. Error bars correspond to 95%-confidence intervals and account for two-way clustering by player and game.

The right panel of Figure 8 repeats this analysis for L-moves. Choice frequencies for L-moves are much smaller and less precisely estimated. This is partly due to the relatively low number of choice sets with more than three L- alternatives (cf. Figure 4). Comparing mean choice frequencies within the same class, we find that out of fifty-six possible pairwise comparisons, over 80% have the correct sign according to Prediction 3 and over 55% are statistically significant at the 5%-level. Only three comparisons, corresponding to about 5% of the total, have the wrong sign and are significant at the 5%-level. Given that some false positives are expected even if the null hypothesis holds, we conclude that the data are broadly consistent with the theory.

Under the maintained assumption of noisy evaluations, the ranking of choice frequencies for W-moves in the left panel of Figure 8 is inconsistent with standard maximization (cf. Proposition 1). This is perhaps not surprising. First, the majority of games in our data are best thought of as "speed chess," and players may simply not have enough time to seriously consider all available moves. Second, some of the choice sets in the data are so large that DMs might find it challenging to evaluate all, or even most, alternatives. Third, the data include choice sets from which skilled-enough players may instantly recognize and gravitate towards very simple moves. For all of these reasons, we view the above findings as a "proof of concept" for our empirical test. It indeed distinguishes satisficing from maximization in an environment where satisficing is most likely to occur. It is entirely possible that choice behavior exhibits different patterns when choice sets are small, time pressure is not an issue, or none of the available moves are simple.

To investigate this possibility, we compute within-choice-set correlations between choice frequencies and moves' relative complexity in such settings. Table 3 presents our estimates, which are based on the following econometric model:

$$Choose_a = \beta Percentile\, Rank_a + \chi_{A,p} + \epsilon_a, \quad (3)$$

where $Choose_a$ denotes an indicator for whether a player p facing choice set A picked move a, $\chi_{A,p}$ is a fixed effect for the particular decision problem of that player (i.e., the task of choosing a move from A), and $Percentile\, Rank_a$ corresponds to the respective move's DTM-based percentile ranking among all available moves of the same type. By construction, this variable is equal to zero for the simplest W- and L-moves in each choice set, and one for the most complex ones.[22]

The coefficient of interest is β. Under the parametric assumptions behind eq. (3), β corresponds to the average difference in choice probability between the simplest and most complex moves of the same type. Since our empirical specification controls for $\chi_{A,p}$, all identifying variation comes from comparing alternatives within the same choice set.

Cols. (1A) and (1B) of Table 3 confirm that, in the full sample, choice frequencies and relative complexity are negatively correlated among W-moves, and positively correlated among L-moves. Taking the point estimates at face value, the simplest winning move in a choice set is about 30 percentage points more likely to be chosen than its most complex counterpart, while the simplest losing move is about 0.3 percentage points less likely to be selected than the most complex one. Relative to the respective mean choice frequencies, the coefficients in cols. (1A) and (1B) are very large.

The next three columns restrict attention to settings that meet one of the following criteria: (i) long time controls (so that each player has, in expectation, at least twenty-five minutes for deliberation per game); (ii) small choice sets (with ten or fewer moves); (iii) none of the available alternatives are easily recognizable as good (because minimal DTM exceeds fifty). Consistent with satisficing, we continue to find that W-moves' choice frequencies are strongly negatively correlated with their complexity. For L-moves, the opposite pattern holds.

The last column presents estimates of β, excluding the simplest move of a given type in every choice set. Although the coefficient in col. (5A) does decrease in magnitude, it remains very large relative to the mean choice frequency of the remaining moves. Thus, Prediction 3

[22] Formally, $Percentile\, Rank_a \equiv \frac{r_a - 1}{M_A - 1}$ where r_a corresponds to move a's ordinal DTM-based rank among moves of the same type in choice set A (simplest = 1, second-simplest= 2, and so on), and M_A denotes the total number of moves of that type. Using moves' percentile rank enables us to sidestep issues related to the fact that different choice sets contain different numbers of W- and L-moves. Appendix Table AT.1 presents results based on ordinal rather than percentile rank. Both sets of results are qualitatively equivalent.

Table 3: Within-Choice-Set Correlations between Relative Complexity and Choice Frequencies

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–30.147 (0.017)	–32.921 (0.116)	–54.851 (0.044)	–23.424 (0.064)	–9.009 (0.007)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	10.305	11.129	26.210	14.992	4.301
R^2	0.234	0.246	0.319	0.208	0.133
N	3,435,257,516	92,878,586	148,599,747	89,336,853	2,986,665,574

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.266 (0.005)	0.148 (0.019)	1.308 (0.030)	0.189 (0.051)	0.576 (0.012)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.452	0.188	1.352	0.588	0.498
R^2	0.277	0.276	0.306	0.344	0.270
N	277,507,532	8,300,536	34,762,492	4,095,758	175,675,645

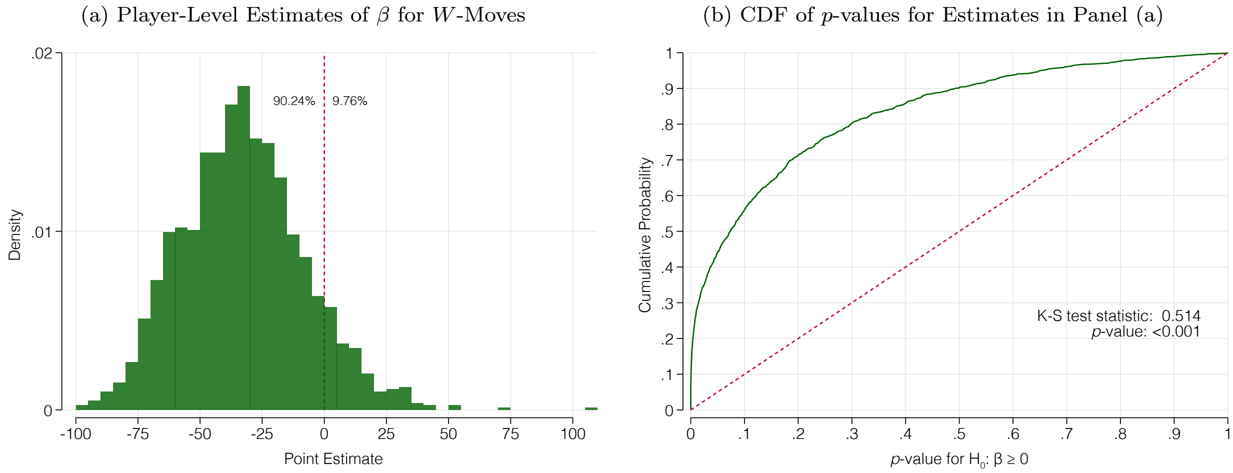
Notes: Entries are coefficients and standard errors from estimating β in eq. (3) by ordinary least squares. The upper panel restricts attention to different samples of W-moves, while the lower panel focuses on L-alternatives. As explained in the text, all regressions control for a decision problem fixed effect, i.e., a fixed effect for a particular player and the set from which she chooses. The unit of observation in each regression is an available legal W- or L-move, and observations are reweighted so that all moves for a given player and all players receive equal total weight. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective move's percentile rank, as defined in the text. Standard errors are two-way clustered by player and game, and are shown in parentheses.

continues to hold.

Going beyond aggregate correlations, Figure 9 examines the relationship between choice frequencies and relative complexity at the level of individual players. To give maximization its best shot, we restrict attention to decision problems that simultaneously meet criteria (i), (ii), and (iii) above. In these settings, players have time to consider all available alternatives, none of which are easily recognizable as good. Focusing on W-moves in this highly selected sample, we estimate the regression model in eq. (3) for every player for whom we observe at least ten such decisions.

The left panel of Figure 9 plots the distribution of the resulting point estimates. For over 90% of players, choice frequencies are negatively correlated with W-moves' relative complexity. The right panel of Figure 9 shows the empirical CDF of the one-sided p-values associated with the estimates on the left. The relevant p-values are one-sided because the null

Figure 9: Testing Noisy Maximization at the Player Level



Notes: Figure presents player-level tests of Proposition 1. Panel (a) plots histograms of estimates of β in eq. (3), restricting attention to W-moves in decision problems in which a player chooses from ten or fewer alternatives with minimal DTM among W-moves of at least fifty, subject to classical or less restrictive time controls. We estimate $\hat{\beta}$ separately for all players for whom we observe at least ten such decision problems. Estimates are scaled to be directly comparable to their counterparts in the upper panel of Table 3. Panel (b) shows the empirical CDF of the one-sided p-values associated with the point estimates in panel (a). It also shows results from a Kolmogorov-Smirnov test against the null hypothesis of a uniform distribution of p-values. All p-values account for clustering across moves in the same choice set.

hypothesis of noisy maximization implies that $\beta \geq 0$ (cf. Proposition 1). Under this null, the distribution of p-values should first-order stochastically dominate the uniform distribution.[23] This, however, is not what we observe. The actual distribution of p-values is first-order stochastically dominated by the uniform distribution and a formal Kolmogorov-Smirnov test rejects the limit case of uniformity at the 99%-confidence level.[24]

We conclude that simpler winning moves tend to be chosen more frequently than more complex ones, even in settings in which players are *a priori* most likely to exhibit maximization-based behavior. We thus reject noisy maximization as the primary choice procedure in the data.

6. Moderators of Object Complexity

Two key factors that may attenuate the adverse impact of object complexity on the quality of choice are deliberation time and the DM's skill, i.e., her cognitive ability, experience, or expertise. Holding deliberation time fixed, a more skillful player is likely capable of examining moves in greater depth, and hence faces less noise in the evaluation process. Similarly, when

[23]That is, we should have $\Pr(p \leq \alpha \mid H_0) \leq \alpha$ for all $\alpha \in [0, 1]$. This claim follows from Definition 8.3.26 and Theorem 8.3.27 in Casella and Berger (2001). The intuition behind it is as follows. If $\beta = 0$, then the observed p-values should be uniformly distributed over the unit interval. If $\beta > 0$, however, then we would expect to see fewer small (one-sided) p-values and more large ones, implying first-order stochastic dominance.

[24]This finding does not depend on the p-values being one-sided. With two-sided p-values a Kolmogorov-Smirnov test would still reject uniformity at the 99%-confidence level.

afforded more time to choose, any given DM might improve the accuracy of her evaluations by developing the subsequent game tree to a greater depth.

One way to incorporate these ideas into the theory is to assume that the variance of the evaluation error in eq. (1) depends not only on an object's inherent complexity but also on the DM's skill S, and the amount of time T that she has to make a decision. In symbols, $\text{Var}(\epsilon_x) = f(\sigma_x, S, T)$. If $\frac{\partial f}{\partial \sigma_x} > 0$, $\frac{\partial f}{\partial S} < 0$, and $\frac{\partial f}{\partial T} < 0$, then skill and time help offset the adverse effect of object complexity on the quality of choice. In the remainder of this section, we empirically test this conjecture. We also examine potential interaction effects between objects' inherent complexity, skill, and deliberation time, i.e., the cross derivatives of f.

In order to explore the role of skill, we distinguish between titled and untitled players. About 1.5% of decisions in our data are made by users who hold real-world chess titles, such as candidate master, national master, international master, or grandmaster. These are some of the best chess players in the world, many of whom are professionals. By contrast, untitled users in our sample are best thought of as experienced hobbyists. It is thus plausible that titled players are, on average, more skilled than untitled ones. Given the sheer size of the data, we observe enough decisions by either type of player to ask whether their mistake frequencies are systematically different, and, if so, when (i.e., how skill and object complexity interact).

We study the first of these questions by estimating the specification

$$(4) \qquad Error_d = \tau Titled_p + \xi_A + \varepsilon_d,$$

where $Error_d$ indicates whether decision d from choice set A by player p is a mistake, and $Titled_p$ indicates whether player p holds a title of candidate master or higher. We control for the difficulty of players' decision problems by adding a fixed effect, ξ_A, for the number of $L-$, D-, and W-moves as well as the vector of DTM values for all L- and W-moves.

Relying on the regression specification above, the results in Table 4 imply that titled players are, on average, significantly less likely to make mistakes than untitled ones. Cols. (1) and (2) show fewer mistakes for all board configurations, while cols. (3) and (4) focus on choice sets that do not contain D-moves. The estimates in cols. (2) and (4) even suggest that grandmasters are less likely to err than other titled players facing similar choice sets. In chess, grandmasters are typically regarded as the "best of the best," but they only account for about 0.15% of decisions in our data, which may explain why the difference between them and other titled players is only statistically significant in col. (2).

To address the question of how skill and object complexity interact, Figure 10 studies whether mistake frequencies for titled and untitled players depend on the minimal DTM among W-moves in the choice set. The left panel of Figure 10 plots raw mistake frequencies

Table 4: Titled vs. Untitled Players' Frequency of Mistakes

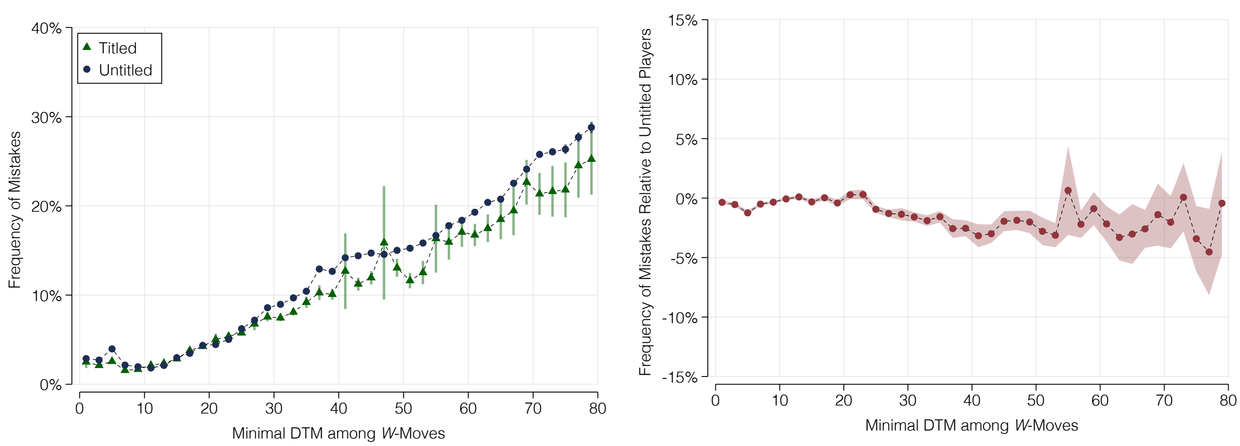
	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–0.813		–0.179	
	(0.048)		(0.061)	
Other Title (÷ 100)		–0.768		–0.155
		(0.051)		(0.064)
Grandmaster (÷ 100)		–1.134		–0.350
		(0.130)		(0.197)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	< 0.001	< 0.001	0.004	0.012
H_0: Grandmasters = Other Titled Players		0.009		0.344
Fixed Effects:				
Number of W- × D- × L-Moves				
× Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	6.039	6.039	2.748	2.748
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.366	0.366	0.410	0.410
N	226,955,095	226,955,095	27,600,514	27,600,514

Notes: Entries are coefficients and standard errors from estimating τ in variants of eq. (4) by ordinary least squares. As explained in the text, all regressions include a fixed effect for the combination of the number of W-, D-, and L-moves and the vector of all DTM values of the W- and L-moves in the choice set. The unit of observation in each regression is a decision problem, and individual observations are reweighted so that all players receive equal total weight. The sample in cols. (1) and (2) includes all board configurations in our data, whereas cols. (3) and (4) restrict attention to configurations for which the associated choice sets do not contain D-moves. All estimates are scaled to correspond to the percentage-point change in the probability of an error associated with a one-unit increase in the respective regressor. Standard errors are two-way clustered by player and game, and are shown in parentheses.

Figure 10: Titled vs. Untitled Players' Frequency of Mistakes, by Minimal DTM among W-Moves

(a) Raw Data

(b) Controlled Difference between Titled and Untitled Players



Notes: Figure compares mistake frequencies for titled and untitled players by minimal DTM among available W-moves. Panel (a) plots raw mistake frequencies, while panel (b) shows point estimates and 95%-confidence intervals for ϕ_i in eq. (5), i.e., the difference in error rates between titled and untitled players when the length of the shortest winning path equals i. Negative point estimates indicate fewer mistakes among titled players. Confidence intervals account for two-way clustering by player and game.

for titled and untitled players. For minimal DTM values below thirty, we observe very small, if any, differences between both sets of players. Only for choice sets in which evaluating W-moves is arguably difficult do we see that titled players commit significantly fewer mistakes. In other words, skill matters, but only when objects' inherent complexity is high.

A potential issue with the raw data is that titled and untitled players might be choosing from sets that are systematically different, even conditional on the same minimal DTM. The right panel of Figure 10 addresses this concern by estimating the difference in the frequency of mistakes controlling for the number of $W-$, D-, and L-moves as well as the DTM values of all W- and L-moves. Specifically, the results in this panel correspond to $\{\phi_i\}$ in the following regression:

$$(5) \qquad Error_d = \sum_{i=1}^{80} \phi_i \lfloor DTM \rfloor_{i,A} \times Titled_p + \xi_A + \varepsilon_d,$$

where $Error_d$, $Titled_p$, and ξ_A are defined as in specification (4) above; and $\lfloor DTM \rfloor_{i,A}$ is an indicator equal to one if the minimal DTM among all W-moves in choice set A is equal to i.

The regression estimates in the right panel accord with the the raw data. Going back to the theory, our findings do not only suggest that skilled players make fewer mistakes, i.e., that $\frac{\partial f}{\partial S} < 0$, but also that skill *moderates* the adverse impact of complexity on the quality of decisions, i.e., $\frac{\partial f}{\partial \sigma_x \partial S} \leq 0$ with strict inequality for large σ.

To explore the implications of deliberation time, we leverage the fact that games on Lichess are subject to various time controls, which govern how fast decisions need to be made. For the sake of parsimony, we broadly categorize the available options as follows: (i) ultra-short time controls (UltraBullet and Bullet), which result in matches that conclude within a few minutes; (ii) short time controls (Blitz and Rapid), under which games have an expected duration of up to fifty minutes; and (iii) long time controls (Classical and Correspondence), resulting in matches that can last hours, or even days.

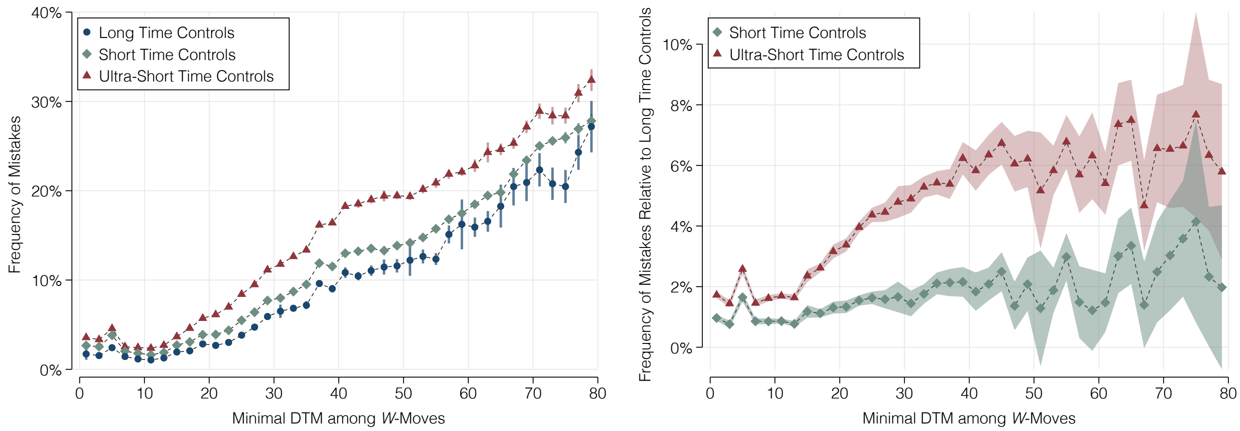
As expected, Figure 11 shows that players are less likely to make mistakes when they are given more time for a decision. Importantly, the gap in the frequency of mistakes across different time controls tends to become wider as the minimal DTM among W-moves increases. This observation is borne out in the raw data (left panel) as well as in regression estimates analogous to those in Figure 10 (right panel). In sum, the evidence in Figures 10 and 11 implies that both skill and time attenuate the adverse impact of complexity on the quality of decisions.

Going one step further, we can ask whether time and skill are complements or substitutes in decision-making. To answer this question, we implement a simple difference-in-differences approach. That is, we compare the difference in the frequency of mistakes between titled

Figure 11: Mistake Frequencies, by Time Controls and Minimal DTM among W-Moves

(a) Mistakes Frequencies in the Raw Data

(b) Controlled Difference Relative to Long Time Controls

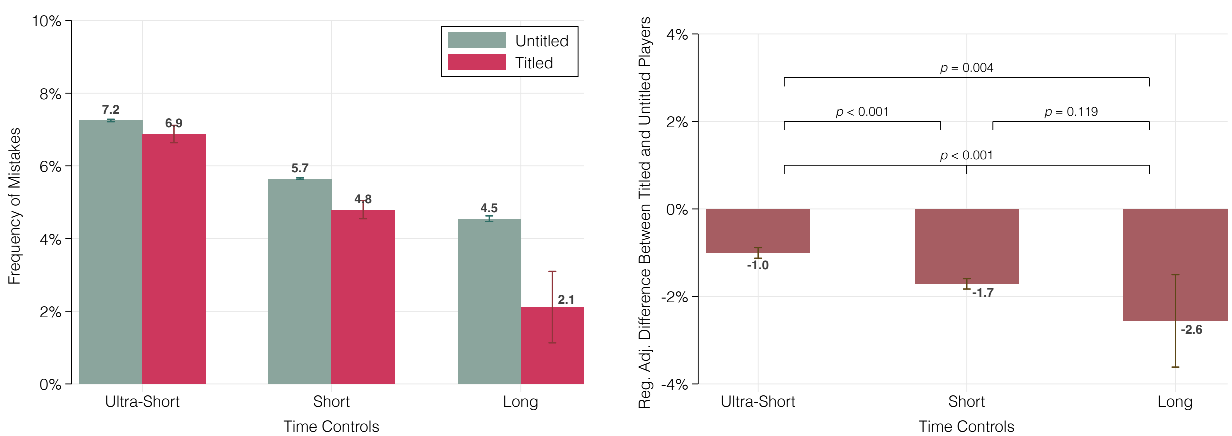


Notes: Figure compares mistake frequencies across different sets of time controls by minimal DTM among available W-moves. Panel (a) plots raw mistake frequencies, while panel (b) shows estimated mean differences relative to long time controls after conditioning on a fixed effect for the number of available moves of each type as well as the vector of DTM values for all L- and W-moves. Confidence intervals account for two-way clustering by player and game.

Figure 12: Mistake Frequencies, by Skill Level and Time Controls

(a) Mistakes Frequencies in the Raw Data

(b) Controlled Difference between Titled and Untitled Players



Notes: Figure compares the frequency of mistakes for titled and untitled players under different time controls. Panel (a) plots raw frequencies, while panel (b) shows estimated mean differences between both sets of players, controlling for the number of available moves of each type as well as the vector of DTM values for all L- and W-moves. Negative values correspond to fewer mistakes among titled players. Moving from the left to the right within each panel, time constraints on players' decision-making become less and less restrictive. Error bars correspond to 95%-confidence intervals and account for two-way clustering by player and game. Braces indicate which estimated differences are being tested for equality.

and untitled players across different sets of time controls. The left panel of Figure 12 shows that, in the raw data, the gap between titled and untitled players widens as both are afforded more time to make decisions. The right panel presents regression adjusted estimates for the difference between both sets of players controlling, as above, for the number and types of all available moves as well as the vector of DTM values. Reassuringly, the numbers in either panel suggest that titled players benefit more from longer time controls than untitled ones. Skill and time thus appear to be complements.

7. Complexity at the Set Level

Moving beyond object complexity and its moderators, there are likely other, set-level factors that complicate choice. For instance, a large literature in psychology argues that choice set size adversely affects decision-making (see, e.g., Iyengar and Lepper 2000; Schwartz 2004; Chernev et al. 2015). In the context of chess, Anderson, Kleinberg, and Mullainathan (2017) show that players are more likely to blunder when choosing from sets that contain a smaller share of W-moves.[25] In this section, we build on their work. We first return to the model and develop its predictions on how the size and composition of the choice set affect mistake probabilities. We then verify these predictions in the data. Finally, we propose a definition of *composition complexity* that may be useful in other applications.

7.1. *Model Predictions*

In the model, every object is characterized by its type and inherent complexity. Our analysis thus far studied object complexity. We now focus on moves' types.

Consider adding a move to the choice set. Random sampling together with noisy evaluations imply that such an addition leads to choice probability shifting away from all previously available moves toward the new one. Thus, if the added alternative is an L- or D-move, the probability of making a mistake increases. If, however, the added move is a W-move then the probability of making a mistake decreases. Comparing the effect of adding a D- to that of adding an L-move, the model predicts that mistake probabilities increase by more when a D-move is added to the choice set.

PREDICTION 4: *For any choice set, adding a D-move to the set increases the probability of a mistake by more than adding an L-move.*

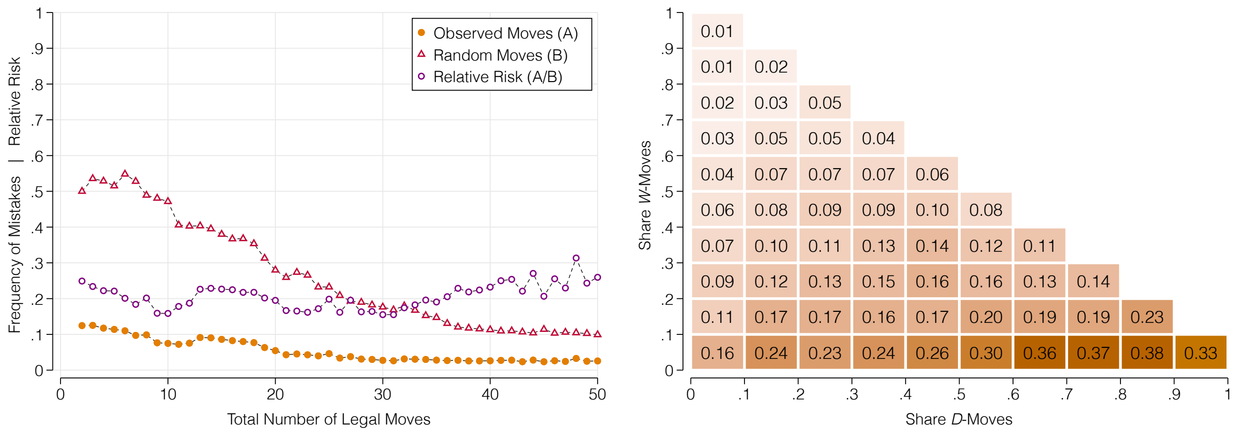
Returning to Figure 1, the intuition for Prediction 4 is that, given any evaluation order, moving from area III to II results in a greater probability of accepting a suboptimal move.

[25] Anderson et al. (2017) do not distinguish between D- and L-alternatives, and they do not consider moves' DTM in any of their analyses.

Figure 13: Mistakes as Function of the Number and Type of Available Moves

(a) Mistakes vs. Choice-Set Size

(b) Mistake Frequencies and Choice-Set Composition



Notes: Panel (a) compares observed mistake rates for choice sets of a particular size with the frequency of mistakes that would be expected if players chose a move at random. Panel (b) shows mistake rates for different shares of W-, D-, and L-moves.

Consider next replacing a W-move with an L- or D-alternative. Because this replacement can be accomplished by first adding an L- or D-move to the set and then removing the W-move, mistake probabilities increase. A more subtle change is to replace an L-move with a D-move. By Prediction 4, this change in the composition of suboptimal alternatives also increases mistake probabilities. Broadly summarizing, our model predicts that what complicates decision-making is not necessarily the size of the choice set but its composition.[26]

7.2. *Empirical Evidence*

Drawing on the raw data, the left panel of Figure 13 depicts the relationship between the size of the choice set and the frequency of mistakes. It also contrasts observed mistake rates with the probability of making a mistake when choosing a move at random. There is a clear negative association between the number of available alternatives and the frequency of mistakes. However, because larger choice sets turn out to contain, on average, a greater share of W-moves, the observed decline in the frequency of mistakes might be mechanical. Measured as a fraction of the baseline risk of making a mistake when choosing at random, error rates remain nearly constant over approximately 85% of our sample. Only for choice sets with more than thirty moves do we observe some increase in relative risk. Taken at face value, the evidence in this panel is difficult to reconcile with elementary notions of choice overload, according to which larger choice sets are *per se* more complex.

[26] The above predictions do not rely on the satisficing component of our model. They would also hold under maximization-based models, provided that the complexity of D-moves is not too small relative to that of L-alternatives.

What appears to affect mistake frequencies is the mix of moves. The right panel of Figure 13 illustrates this point by examining how the rate of mistakes varies with the relative frequency of each type of move in the choice set. The entry on the bottom left, for instance, indicates that when the fraction of W- and D-moves ranges from zero to ten percent each, about 16% of choices end up being mistakes. Comparing rows within the same column, we see mistakes become more frequent as W-moves become relatively more scarce. And reading across columns of the same row, mistakes also tend to increase as D-moves become more frequent. In the bottom row, for instance, the frequency of mistakes increases from 16% to more than 30%. Since comparisons within the same row hold the fraction of W-moves—and hence the mechanical probability of making a mistake—approximately fixed, the raw data suggest the quality of decision-making correlates with the *composition* of the remaining, inferior alternatives.

To more rigorously establish that different types of changes to the choice set affect mistake frequencies as predicted by the theory, we estimate η in the following econometric model:

$$Error_d = \eta Number\, Moves^t_A + \kappa^{\lfloor DTM \rfloor}_A + \theta^{-t}_A + \mu_p + \varepsilon_d. \tag{6}$$

In this specification, $Error_d$ is an indicator for whether decision d of player p choosing from choice set A is a mistake, and $Number\, Moves^t_A$ denotes the number of available legal moves of type $t \in \{W, D, L\}$. Although not required by the theory, we account (in a low-dimensional way) for the contribution of object complexity to mistake frequencies by including $\kappa^{\lfloor DTM \rfloor}_A$, a fixed effect for the minimal DTM among all available W-moves.[27] The fixed effect θ^{-t}_A controls in different ways for the number of moves in the same choice set that are not of type t. As we vary θ^{-t}_A across columns in Table 5, the interpretation of η changes to closely align with the different changes to choice-set composition that we discussed above. For example, col. (6) conditions on the choice-set size and the number of W-moves, so that $\hat{\eta}$ should be interpreted as the change in mistake probabilities when an L-move is replaced with one of type D. By including μ_p, the model in eq. (6) also controls for player fixed effects.

The regression in col. (1) of Table 5 conditions on the number of D- and L-moves. The point estimate in this column therefore indicates that the frequency of mistakes decreases, on average, by .279 p.p. for every W-move that is added to the choice set. This corresponds to almost 5% of the mean frequency of mistakes. In analogous fashion, the coefficients in cols. (2) and (3) imply that mistake frequencies respectively increase by .810 and .247 p.p. as D-

[27] In Appendix Table AT.2, we show that controlling for object complexity is not necessary to obtain results that are substantively equivalent to those in Table 5.

Table 5: Mistakes as a Function of Choice-Set Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of Mistake						
Number of W-Moves ($\div$ 100)	–0.279			–1.242	–0.988		
	(0.001)			(0.002)	(0.003)		
Number of D-Moves ($\div$ 100)		0.810				0.841	
		(0.002)				(0.003)	
Number of L-Moves ($\div$ 100)			0.247				
			(0.003)				
Total Number of Moves ($\div$ 100)							–0.017
							(0.001)
Fixed Effects:							
Number of D- $\times$ L-Moves	Yes	No	No	No	No	No	No
Number of W- $\times$ L-Moves	No	Yes	No	No	No	No	No
Number of W- $\times$ D-Moves	No	No	Yes	No	No	No	No
Total Moves $\times$ Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves $\times$ Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves $\times$ Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- $\times$ D- $\times$ L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.119	0.144	0.153	0.120	0.143	0.151	0.155
N	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095

Notes: Entries are coefficients and standard errors from estimating η in variants of eq. (6) by ordinary least squares. All regressions control for player fixed effects as well as the minimal DTM among W-moves in the same choice set. Controls for the number and types of other moves vary across columns. The unit of observation in each regression is a decision, and individual observations are reweighted so that all players receive equal total weight. The sample in every column includes all board configurations in our data. All estimates are scaled to correspond to the percentage point change in the probability of an error associated with a one-unit increase in the respective regressor. Standard errors are two-way clustered by player and game, and are shown in parentheses.

and L-moves are added to the choice set. The evidence in the first three columns of Table 5 thus shows that whether mistakes become more or less frequent as the choice set expands depends on the type of move that is being added.

This pattern helps to explain why the raw data yield little evidence of choice overload, i.e., why mistake frequencies are not uniformly higher for larger choice sets. Perhaps more importantly, the increase in the frequency of mistakes is significantly greater when a D- rather than an L-move is added to the choice set ($p < .001$), consistent with Prediction 4.

Changes to the choice set that involve replacing one move with another one of a different type are considered in cols. (4)–(6). The coefficient in col. (4) refers to a W-move being replaced by a D-alternative, whereas that in col. (5) refers to a switch between W- and L-moves. The point estimate in col. (6) tells us that, on average, mistake frequencies increase by .841 p.p.—or about 15% of the mean mistake frequency—when an L-move is replaced by a D-alternative.

Col. (7) considers homothetic enlargements of the choice set, i.e., expansions that hold the *share* of W-, D-, and L-moves fixed. After carefully controlling for the composition of available alternatives, we find no evidence to suggest that larger choice sets are associated with more mistakes. Not only is the coefficient in col. (7) more than an order of magnitude smaller than the other estimates in Table 5, it is, if anything, negative and very precisely estimated.

In sum, the results in Table 5 are consistent with the theory's predictions. Moreover, the point estimates in cols. (1)–(6) are nontrivial in size—ranging from 4% to 21% of the mean mistake rate in the data. Our findings, therefore, imply that complexity at the set level is empirically relevant and closely related to the composition rather than the size of the choice set.

7.3. *Set-Level Complexity Beyond Chess*

Motivated by these findings, we propose a concise, portable definition of what we call composition complexity. Let A be a choice set, with $V(A) = \{v_x \mid x \in A\}$ denoting the set of values in A. Let $\bar{v}(A)$ be the largest value in $V(A)$, and let $N(v, A)$ be the number of objects in A that have value v. With this notation in hand and assuming that $|V(A)| > 1$, the composition complexity of A is:

$$\Gamma(A) \equiv -N(\bar{v}(A), A) + \sum_{v \in V(A) \setminus \bar{v}(A)} \alpha(\bar{v}(A) - v) N(v, A), \tag{7}$$

where $\alpha(\cdot)$ is a positive and monotonically decreasing function.

The first term in this expression implies that as the number of maximal objects in the choice set increases, the choice task becomes easier. The function $\alpha(\cdot)$ in the second term captures the key features of the analysis above. Positivity of $\alpha(\cdot)$ means that adding non-maximal objects increases complexity, and monotonicity implies that complexity increases as suboptimal objects become more similar to optimal ones.[28]

8. Concluding Remarks

This paper studies how complexity at the level of individual objects affects decision-making. We establish that object complexity affects choice in systematic ways. Choice frequencies of optimal alternatives decrease in the respective object's complexity, whereas choice frequencies of suboptimal alternatives follow the opposite pattern. We also develop an empirical test that leverages complexity to distinguish between satisficing and a large class of maximization-based models. Based on this test, we reject maximization as the primary choice procedure in our data.

Taking a step back, our analysis hints at the potential of developing a ranking of choice sets according to their complexity. Take two sets A and B with the same support of values. Based on our results, set B is more complex than set A if it is possible to transition from A to B by iteratively applying any of the following operations:

- increasing the object complexity of some alternative,
- adding a suboptimal alternative, or
- removing an optimal one.

The resulting binary relation is transitive but incomplete, suggesting an avenue for future research.

More broadly, our results on complexity in decision-making offer insights that may be important in other real-world settings. For example, if consumers' choice sets include products by different firms, then firms with superior offerings may wish to simplify their product descriptions. In contrast, firms with inferior offerings have an incentive to make their products more complex. Firms also have an incentive to affect the order in which satisficing consumers consider offerings even if doing so is costly. Similar incentives may be present in the design and presentation of policy proposals, in strategic games, and in various other settings.

[28] The exact cardinal specification of $\alpha(\cdot)$ determines, together with the composition of the choice set, how composition complexity changes when an object with value above $\bar{v}(A)$ is added to A.

References

Abeler, J. and S. Jäger (2015). Complex Tax Incentives. *American Economic Journal: Economic Policy 7*(3), 1–28.

Abreu, D. and A. Rubinstein (1988). The Structure of Nash Equilibrium in Repeated Games with Finite Automata. *Econometrica 56*(6), 1259–1281.

Anderson, A., J. Kleinberg, and S. Mullainathan (2017). Assessing Human Error Against a Benchmark of Perfection. *ACM Transactions on Knowledge Discovery from Data (TKDD) 11*(4), 45.

Bossaerts, P. and C. Murawski (2017). Computational Complexity and Human Decision-Making. *Trends in Cognitive Sciences 21*(12), 917–929.

Camerer, C. (2003). *Behavioral Game Theory.* Princeton, NJ: Princeton University Press.

Caplin, A., M. Dean, and D. Martin (2011). Search and Satisficing. *American Economic Review 101*(7), 2899–2922.

Casella, G. and R. Berger (2001). *Statistical Inference* (2nd ed.). Boston: Cenage.

Chernev, A., U. Böckenholt, and J. Goodman (2015). Choice Overload: A Conceptual Review and Meta-Analysis. *Journal of Consumer Psychology 25*(2), 333–358.

Chiappori, P.-A., S. Levitt, and T. Groseclose (2002). Testing Mixed-Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer. *American Economic Review 92*(4), 1138–1151.

Enke, B. and T. Graeber (2021). Cognitive Uncertainty. NBER Working Paper No. 26518.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006). Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review 96*(4), 1043–1068.

Glickman, M. E. (1999). Parameter Estimation in Large Dynamic Paired Comparison Experiments. *Journal of the Royal Statistical Society: Series C 48*(3), 377–394.

Glickman, M. E. (2000). The Glicko-2 System for Rating Players in Head-to-Head Competition. mimeographed, Boston University.

Hsu, S.-H., C.-Y. Huang, and C.-T. Tang (2007). Minimax Play at Wimbledon: Comment. *American Economic Review 97*(1), 517–523.

Huck, S. and G. Weizsäcker (1999). Risk, Complexity, and Deviations from Expected-Value Maximization: Results of a Lottery Choice Experiment. *Journal of Economic Psychology 20*(6), 699–715.

Iyengar, S. and M. Lepper (2000). When Choice is Demotivating: Can One Desire Too Much of a Good Thing? *Journal of Personality and Social Psychology 79*(6), 995–1006.

Jakobsen, A. M. (2020). A Model of Complex Contracts. *American Economic Review 110*(5), 1243–1273.

Kalai, E. and W. Stanford (1988). Finite Rationality and Interpersonal Complexity in Repeated Games. *Econometrica 56*(2), 397–410.

Levitt, S. and J. List (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal

About the Real World? *Journal of Economic Perspectives 21*(2), 153–174.

Levitt, S., J. List, and D. Reiley (2010). What Happens in the Field Stays in the Field: Exploring Whether Professionals Play Minimax in Laboratory Experiments. *Econometrica 78*(4), 1413–1434.

Levitt, S., J. List, and S. Sadoff (2011). Checkmate: Exploring Backward Induction among Chess Players. *American Economic Review 101*(2), 975–990.

Luce, D. (1959). *Individual Choice Behavior*. New York: Wiley.

Man, R. d. (2013). Syzygy tablebases. available at https://github.com/syzygy1/tb.

Marschak, J. (1960). Binary-Choice Constraints and Random Utility Indicators. In K. Arrow, S. Karlin, and P. Suppes (Eds.), *Stanford Symposium on Mathematical Methods in the Social Sciences*, pp. 312–329. Stanford, CA: Stanford University Press.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zambreka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

Nalimov, E., G. Haworth, and E. Heinz (2000). Space-Efficient Indexing of Chess Endgame Tables. *ICGA Journal 23*(3), 148–162.

Neumann, J. v. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen 100*(1), 295–320.

Neumann, J. v. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Neyman, A. (1985). Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma. *Economics Letters 19*(3), 227–229.

Oprea, R. (2020). What Makes a Rule Complex? *American Economic Review 110*(12), 3913–3951.

Osborne, M. and A. Rubinstein (1994). *A Course in Game Theory*. Cambridge, MA: MIT Press.

Palacios-Huerta, I. (2003). Professionals Play Minimax. *Review of Economic Studies 70*(2), 395–415.

Palacios-Huerta, I. and O. Volij (2008). Experientia Docet: Professionals Play Minimax in Laboratory Experiments. *Econometrica 76*(1), 71–115.

Palacios-Huerta, I. and O. Volij (2009). Field Centipedes. *American Economic Review 99*(4), 1619–1635.

Rubinstein, A. (1986). Finite Automata Play the Repeated Prisoner's Dilemma. *Journal of Economic Theory 39*(1), 83–96.

Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *Economic Journal 117*(523), 1243–1259.

Rubinstein, A. (2016). A Typology of Players: Between Instinctive and Contemplative. *Quarterly Journal of Economics 131*(2), 859–890.

Salant, Y. (2011). Procedural Analysis of Choice Rules with Applications to Bounded Rationality. *American Economic Review 101*(2), 724–748.

Schwartz, B. (2004). *The Paradox of Choice: Why More Is Less*. New York: HarperCollins.

Simon, H. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics 69*(1), 99–118.

Simon, H. (1972). Theories of Bounded Rationality. In C. McGuire and R. Radner (Eds.), *Decision and Organization*, pp. 161–176. Amsterdam: North-Holland.

Spenkuch, J., P. Montagnes, and D. Magleby (2018). Backward Induction in the Wild? Evidence from Sequential Voting in the US Senate. *American Economic Review 108*(7), 1971–2013.

Thompson, K. (1986). Retrograde Analysis of Certain Endgames. *ICGA Journal 9*(3), 131–139.

Walker, M. and J. Wooders (2001). Minimax Play at Wimbledon. *American Economic Review 91*(5), 1521–1538.

Wilson, A. (2014). Bounded Memory and Biases in Information Processing. *Econometrica 82*(6), 2257–2294.

Wooders, J. (2010). Does Experience Teach? Professionals and Minimax Play in the Lab. *Econometrica 78*(3), 1143–1154.

Woodford, M. (2020). Modeling Imprecision in Perception, Valuation, and Choice. *Annual Review of Economics 12*, 579–601.

Zermelo, E. (1913). Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. In E. Hobson and A. Love (Eds.), *Proceedings of the Fifth International Congress of Mathematicians (Cambridge 1912)*, pp. 501–504. Cambridge, UK: Cambridge University Press.

Online Appendix

Contents

List of Tables

Appendix A: Proofs

A.1. *Proof of Prediction 1*

By Assumption 1, the choice probability of any W-move is larger than $1/2$, and the choice probability of any D- or L-move is smaller than $1/2$. It thus remains to be proven that the choice probability of any D-move is larger than that of any L-move. To do so, we first state the following lemma.

Lemma 1: *Let $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$. Then,*

$$Pr(X_1 \geq T) > Pr(X_2 \geq T) \iff T(\sigma_2 - \sigma_1) < \mu_1\sigma_2 - \mu_2\sigma_1.$$

Proof: Let Φ denote the CDF of the standard normal distribution. Then,

$$Pr(X_i \geq T) = Pr\left(\frac{X_i - \mu_i}{\sigma_i} \geq \frac{T - \mu_i}{\sigma_i}\right) = 1 - \Phi\left(\frac{T - \mu_i}{\sigma_i}\right).$$

The required ranking of the probabilities holds if and only if

$$\Phi\left(\frac{T - \mu_2}{\sigma_2}\right) > \Phi\left(\frac{T - \mu_1}{\sigma_1}\right) \iff \frac{T - \mu_2}{\sigma_2} > \frac{T - \mu_1}{\sigma_1} \iff T(\sigma_2 - \sigma_1) < \mu_1\sigma_2 - \mu_2\sigma_1.$$

Q.E.D.

The condition $T(\sigma_2 - \sigma_1) < \mu_1\sigma_2 - \mu_2\sigma_1$ identified in Lemma 1 holds when the first random variable corresponds to a D-move x and the second corresponds to an L-move y, because in this case $\mu_1 = 0$, $\mu_2 = L = -W$, $W > T$ by Assumption 1, and $\sigma_1 \geq \frac{1}{2}\sigma_2$ by Assumption 2. Thus, $Pr(u_x \geq T) > Pr(u_y \geq T)$.

To complete the proof, fix a choice set and two evaluation orders of moves O_1 and O_2 that are identical except that x appears before y in O_1, and their locations are switched in O_2. The choice probability of x from these two orderings is

$$Pr(u_x \geq T) \times \Big(Pr(T \text{ not exceeded prior to } x \text{ in } O_1) + Pr(T \text{ not exceeded prior to } x \text{ in } O_2)\Big).$$

The choice probability of y is identical except for $Pr(u_y \geq T)$ replacing $Pr(u_x \geq T)$. The choice probability of x is larger than that of y because

(i) $Pr(u_x \geq T) > Pr(u_y \geq T)$ as we proved above,

(ii) $Pr(T$ not exceeded prior to x in $O_1) = Pr(T$ not exceeded prior to$y$ in $O_2)$ because the evaluation order prior to reaching the first of the two moves x and y is identical in O_1 and O_2, and

(iii) $Pr(T$ not exceeded prior to x in $O_2) > Pr(T$ not exceeded prior to y in $O_1)$ because the order prior to reaching the second of the two moves x and y is identical expect for x appearing in O_1 and being chosen with larger probability than y in O_2.

Since this ranking of the probabilities holds for every pair of evaluation orders in which the locations of x and y are switched, it holds for their choice probabilities from the choice set. *Q.E.D.*

A.2. *Proof of Prediction 2*

We prove the result for W-moves. The proof for D- and L-moves is analogous.

Fix the order in which the DM evaluates moves. Let x denote a W-move. If x is last in the order, its choice probability conditional on reaching it is 1 independently of its complexity σ_x. Otherwise, its choice probability conditional on the order is

$$Pr(T \text{ not exceeded prior to } x) \times Pr(u_x \geq T).$$

The first component in this expression is independent of σ_x, and the second decreases in σ_x because $T < W$. Consequently, the choice probability of any move that appears after x in the order increases in σ_x. Because this holds for any order in which x does not appear last, the result follows. *Q.E.D.*

A.3. *Proof of Prediction 3*

We prove the result for W-moves. The proof for D- and L-moves is analogous.

Assume to the contrary that the choice probability of x is weakly smaller than the choice probability of y. Now increase σ_x until it is equal to σ_y. By Prediction 2, the choice probability of x decreases and the choice probability of y increases, implying the choice probability of x remains smaller than that of y. This is in contrast to the fact that two W-moves with the same object complexity should be chosen with the same probability. *Q.E.D.*

A.4. *Proof of Proposition 1*

Because posterior expected values are ranked in the same way as the estimates, and because the DM chooses the maximal expected value, move x is chosen from some choice set if and only if its realized u-value is larger than the realized u-values of all other moves in the choice set. Thus, to prove the result, it suffices to establish the following:

LEMMA 2: *Let $X_1, \ldots, X_N$ be N independently distributed normal random variables where $X_i \sim N(\mu_i, \sigma_i)$, $\mu_1 = \mu_2$, and $\sigma_1 < \sigma_2$. Then,*

$$P_2 = \mathbb{P}\left[X_2 > \{X_j\}_{j\neq 2}\right] > \mathbb{P}\left[X_1 > \{X_j\}_{j\neq 1}\right] = P_1.$$

PROOF: Without loss of generality, we rescale all RVs so that each X_i is replaced by $\frac{X_i - \mu_1}{\sigma_1}$. After rescaling, we have that $X_1 \sim N(0,1)$ and $X_2 \sim N(0,\sigma)$ where $\sigma = \sigma_2/\sigma_1 > 1$.

Let $h(x)$ denote the probability that $X_3, \ldots, X_N \leq x$. Then, $h(x)$ strictly increases in x. Let F and f (G and g) denote the CDF and PDF of X_1 (X_2) respectively. Then,

$$P_2 - P_1 = \int_{-\infty}^{\infty} h(x)\Big(F(x)g(x) - G(x)f(x)\Big)dx.$$

Because $X_2 - X_1$ is distributed normal with mean 0, we have that

$$0 = \mathbb{P}(X_2 > X_1) - \mathbb{P}(X_1 > X_2) = \int_{-\infty}^{\infty} \Big(F(x)g(x) - G(x)f(x) \Big) dx.$$

Because $h(x)$ strictly increases in x, if we were to show that the function $m(x) = F(x)g(x) - G(x)f(x)$ crosses 0 exactly once from below at some $\hat{x}$ then the conclusion of the lemma would follow because in this case,

$$\begin{aligned}
\int_{-\infty}^{\infty} h(x) \Big(F(x)g(x) - G(x)f(x) \Big) dx &= \int_{-\infty}^{\hat{x}} h(x)m(x)dx + \int_{\hat{x}}^{\infty} h(x)m(x)dx \\
&> \int_{-\infty}^{\hat{x}} h(\hat{x})m(x)dx + \int_{\hat{x}}^{\infty} h(\hat{x})m(x)dx \\
&= h(\hat{x}) \int_{-\infty}^{\infty} m(x)dx = 0.
\end{aligned}$$

To conclude the proof, it thus suffices to show that $m(x)$ crosses 0 exactly once from below at $\hat{x} > 0$. We do so by showing that

(i) $m(x) \leq 0$ for all $x \leq 0$ with strict inequality for $x = 0$,

(ii) $m(x) > 0$ for some $x > 0$, and

(iii) $m(x)$ cannot cross 0 from above at some $x > 0$.

To verify (i), observe first that $m(0) = 1/2(g(0) - f(0)) < 0$. To prove that $m(x) \leq 0$ for any $x < 0$, fix $x < 0$ and consider the interval $[a, x]$ with $a < x$. By Cauchy's mean value theorem, there exists $a < c < x$ such that

$$(F(x) - F(a))g(c) = (G(x) - G(a))f(c).$$

Since the ratio $g(x)/f(x)$ decreases in x for $x < 0$, we have that

$$(F(x) - F(a))g(x) < (G(x) - G(a))f(x),$$

which is true for any $a < x$. We take a to $-\infty$ to obtain $F(x)g(x) \leq G(x)f(x)$, i.e., $m(x) \leq 0$.

To verify (ii), observe that at the point of intersection x between f and g to the right of 0, $m(x)$ is proportional to $F(x) - G(x)$ and $F(x) - G(x) > 0$.

To verify (iii), assume to the contrary that $m(x)$ crosses 0 from above at some $x > 0$. Then, the following two equations should hold at x:

(1) $F(x)g(x) - G(x)f(x) = 0$, and

(2) $m'(x) \leq 0$ where m' is the derivative of m.

We develop $m'(x)$ to show (1) and (2) cannot hold simultaneously. Because the PDF z of a normal distribution with mean 0 and variance σ^2 satisfies the identity $z'(x) = -\frac{xz(x)}{\sigma^2}$, we have that

$$0 \geq m'(x) = f(x)g(x) + F(x)g'(x) - f(x)g(x) - G(x)f'(x) = xf(x)G(x) - \frac{xg(x)}{\sigma^2}F(x),$$

implying $f(x)G(x) \leq \frac{g(x)}{\sigma^2}F(x) < g(x)F(x)$ in contradiction to (1). *Q.E.D.*

A.5. *Proof of Prediction 4*

By Assumptions 1 and 2 and Lemma 1, we have that $Pr(u_x \geq T) > Pr(u_y \geq T)$ for a D-move x and an L-move y.

Fix a choice set A. Let $A_x = A \cup \{x\}$ and $A_y = A \cup \{y\}$. Fix an evaluation order O_1 of A_x in which x does not appear last, and an evaluation order O_2 of A_y, which is identical to O_1 except that y replaces x. The probability of making a mistake prior to x in O_1 and y in O_2 is identical. The conditional probability of making a mistake when evaluating x in O_1 is larger than when evaluating y in O_2. The probability of making a mistake conditional on $x < T$ in O_1 is identical to the corresponding conditional probability in O_2. The result follows. *Q.E.D.*

Appendix B: Data Appendix

Our data on endgame moves come from `lichess.org`. Every month, Lichess releases database extracts covering all rated chess games between two human players that were hosted on its platform during the previous month. These extracts are made available in the human-readable PGN format at `https://database.lichess.org`, and include basic facts about each game (including players' usernames and ratings, date and time of the game, time controls, ultimate outcome, etc.), the exact sequence of moves, as well as, starting April 2017, the clock reading at the end of each move.

We downloaded and processed all extracts through August 2020, filtering on endgame positions with six or fewer pieces. We then spent about 600,000 CPU-hours querying the Nalimov and Syzygy endgame tablebases for information on depth to mate (DTM) and the type of each available legal move (i.e., W, D, or L) in these positions. The 6-men Syzygy and Nalimov endgame databases are available at `http://tablebase.sesse.net` (Syzygy: 150GB; Nalimov: 1.2TB). Because Syzygy tablebases take into account the 50-move rule, we rely on them to determine the type of each move, whereas information on DTM comes from Nalimov's database. The only board configurations with six or fewer pieces that are not covered in the latter are *(i)* ones in which a lone king faces five other pieces, and *(ii)* positions with castling rights. The former are generally uninteresting because 98.8% of available legal moves are of type W, and the latter are extremely rare in the Lichess data (< .01% of moves in nontrivial endgame positions).

The sample for our main analysis restricts attention to decision problems in *(i)* board positions with six or fewer pieces with *(ii)* available information on the types of all available legal moves and the DTM of all available W- and L-moves, in which *(iii)* there are one or more legal W-moves and at least one D- or L-alternative, *(iv)* excluding the first 1,000 such decision problems for every user.

Appendix C: Robustness Checks

C.1. *Controlling for Time Left*

Since the timing of decisions is endogenous, we do not control for it in our main analysis. We do, however, obtain qualitatively equivalent findings when we account for it. To show this, we replicate the tables in the main text. In Appendix Tables AT.3–AT.6, we control for the time that remains on the player's clock when it is her turn to move. In Tables AT.7–AT.10, we control for the time that was left per move if the player were to follow the shortest W-path. Regardless of how we account for the possibility that players face time pressure, our findings on how complexity affects decision-making are qualitatively equivalent.

C.2. *Restricting Attention to Games with Long Time Controls*

In Appendix Tables AT.11–AT.14, we replicate our main results restricting attention to games played under "classical" and "correspondence" time controls. The former typically allow more than 25 minutes of deliberation per side, whereas the latter usually take days or weeks to complete. More specifically, Lichess classifies the time controls in a game as classical if and only if the estimated time per side exceeds 1,500 seconds, with the estimated time per side: (initial clock time) + 40 $\times$ (clock increment). In correspondence games, the time limit is measured in days per move. The results with long time controls are qualitatively equivalent to those in the main text.

C.3. *Restricting Attention to Board Positions with High Minimal DTM*

In Appendix Tables AT.15–AT.18, we replicate the results in the main text, restricting attention to board positions in which the minimal DTM among W-moves exceeds fifty. These are positions in which it is *a priori* highly unlikely that players can recognize either the type of a move or its DTM without careful evaluation, as assumed in our model. Reassuringly, the results from this smaller sample are qualitatively equivalent to those in the main text.

C.4. *Weighting Observations Equally*

Our findings are similar when we do not reweight individual observations so that all players receive equal weight. To show this, we replicate the tables in the main text, weighting all observations equally. The respective results are shown in Appendix Tables AT.19–AT.22.

Appendix D: Replication with Independent Data from *The Week in Chess*

Appendix Tables AT.23–AT.26 replicate the tables in the main text, using an independent dataset that we obtained from *The Week in Chess* (TWIC). TWIC is a free, weekly publication that "rounds up the most important chess" games from the previous week (see `https://theweekinchess.com`). Most of these games are played between elite players in national and international tournaments, or chess leagues.

Our data include all games covered in TWIC between September 1994 and May 2020. In total, we observe 536,674 decision problems in endgame positions with six or fewer pieces, one or more legal W- and at least one D- or L-moves. The choice sets in these decision problems contain 9,067,040 legal moves.

Besides being more than two orders of magnitude smaller, the most important difference between the TWIC and Lichess data is that the former admit much less variation in players' skill. Chess players in high-profile tournaments tend to be better than the average experienced player on Lichess. This fact is reflected in a significantly lower frequency of mistakes in the TWIC data. Nonetheless, the comparative statics in Appendix Tables AT.23–AT.26 are similar to those in the main text.

Appendix Tables

Appendix Table AT.1: Replication of Table 3, Using Ordinal Rank

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Ordinal Rank (÷ 100)	–0.862 (0.001)	–1.036 (0.004)	–12.561 (0.009)	–1.803 (0.004)	–0.327 (0.000)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	10.305	11.129	26.210	14.992	4.301
R^2	0.166	0.171	0.247	0.179	0.127
N	3,435,257,516	92,878,586	148,599,747	89,336,853	2,986,665,574

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Ordinal Rank (÷ 100)	0.071 (0.001)	0.042 (0.004)	0.470 (0.009)	0.042 (0.006)	0.079 (0.001)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.452	0.188	1.352	0.588	0.498
R^2	0.277	0.276	0.306	0.344	0.271
N	277,507,532	8,300,536	34,762,492	4,095,758	175,675,645

Notes: See Table 3 in the main text. The only difference between this table and that in the text is that the results above are rely on moves' ordinal rather than percentile rank.

Appendix Table AT.2: Replication of Table 5, Without Controlling for Object Complexity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Probability of Mistake			
Number of W-Moves (÷ 100)	–0.307			–1.344	–1.173		
	(0.001)			(0.002)	(0.003)		
Number of D-Moves (÷ 100)		0.819				0.766	
		(0.002)				(0.003)	
Number of L-Moves (÷ 100)			0.361				
			(0.003)				
Total Number of Moves (÷ 100)							–0.039
							(0.001)
Fixed Effects:							
Number of D- × L-Moves	Yes	No	No	No	No	No	No
Number of W- × L-Moves	No	Yes	No	No	No	No	No
Number of W- × D-Moves	No	No	Yes	No	No	No	No
Total Moves × Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves × Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves × Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- × D- × L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.101	0.127	0.139	0.103	0.128	0.136	0.141
N	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above do not control for the minimal DTM among W-moves.

Appendix Table AT.3: Replication of Table 2, Controlling for Time Left on Clock

	(1)	(2)	(3)	(4)	(5)	(6)
			Probability of Choosing Move			
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM × W-Move (÷ 100)	–1.140		–0.975	–0.518		–0.332
	(0.002)		(0.002)	(0.004)		(0.003)
DTM × L-Move (÷ 100)		0.034	0.007		0.066	0.055
		(0.001)	(0.001)		(0.002)	(0.003)
W-Move (÷ 100)			63.533			41.926
			(0.071)			(0.118)
Fixed Effects:						
Number of W- × D- × L-Moves × Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	9.397	0.785	8.495	10.881	1.183	8.078
R^2	0.372	0.211	0.368	0.513	0.212	0.494
N	3,238,254,715	276,991,030	3,515,245,745	372,397,046	104,556,541	476,953,587

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that the results above also control for the clock reading when it is a player's turn to move.

Appendix Table AT.4: Replication of Table 3, Controlling for Time Left on Clock

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–30.105 (0.018)	–32.975 (0.150)	–54.813 (0.045)	–23.409 (0.065)	–9.002 (0.007)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	10.305	11.129	26.210	14.992	4.301
R^2	0.234	0.247	0.319	0.208	0.133
N	3,217,115,234	56,772,033	138,639,890	83,387,752	2,797,083,668

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.267 (0.005)	0.149 (0.025)	1.318 (0.031)	0.190 (0.053)	0.577 (0.012)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.452	0.188	1.352	0.588	0.498
R^2	0.277	0.279	0.306	0.344	0.271
N	259,240,947	5,014,246	32,452,348	3,829,648	164,138,685

Notes: See Table 3 in the main text. Note that conditional on the included fixed effects there is no independent variation in the clock reading when it is a player's turn to move.

Appendix Table AT.5: Replication of Table 4, Controlling for Time Left on Clock

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–0.985 (0.049)		–0.339 (0.063)	
Other Title (÷ 100)		–0.937 (0.053)		–0.312 (0.065)
Grandmaster (÷ 100)		–1.318 (0.131)		–0.531 (0.200)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	$<$ 0.001	$<$ 0.001	$<$ 0.001	$<$ 0.001
H_0: Grandmasters = Other Titled Players		0.007		0.294
Fixed Effects: Number of W- × D- × L-Moves × Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	6.039	6.039	2.748	2.748
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.367	0.367	0.412	0.412
N	212,295,223	212,295,223	25,771,678	25,771,678

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that the results above also control for the clock reading when it is a player's turn to move.

Appendix Table AT.6: Replication of Table 5, Controlling for Time Left on Clock

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of Mistake						
Number of W-Moves (÷ 100)	–0.281			–1.245	–0.990		
	(0.001)			(0.002)	(0.003)		
Number of D-Moves (÷ 100)		0.811				0.842	
		(0.002)				(0.003)	
Number of L-Moves (÷ 100)			0.247				
			(0.003)				
Total Number of Moves (÷ 100)							–0.018
							(0.001)
Fixed Effects:							
Number of D- × L-Moves	Yes	No	No	No	No	No	No
Number of W- × L-Moves	No	Yes	No	No	No	No	No
Number of W- × D-Moves	No	No	Yes	No	No	No	No
Total Moves × Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves × Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves × Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- × D- × L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.120	0.144	0.154	0.121	0.143	0.152	0.156
N	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above also control for the clock reading when it is a player's turn to move.

Appendix Table AT.7: Replication of Table 2, Controlling for Time Left per Move

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of Choosing Move					
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM × W-Move (÷ 100)	–1.140		–0.975	–0.518		–0.332
	(0.002)		(0.002)	(0.004)		(0.003)
DTM × L-Move (÷ 100)		0.034	0.007		0.066	0.055
		(0.001)	(0.001)		(0.002)	(0.003)
W-Move (÷ 100)			63.532			41.925
			(0.071)			(0.118)
Fixed Effects:						
Number of W- × D- × L-Moves × Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	9.397	0.785	8.495	10.881	1.183	8.078
R^2	0.372	0.211	0.368	0.513	0.211	0.494
N	3,238,254,715	276,991,030	3,515,245,745	372,397,046	104,556,541	476,953,587

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that the results above also control for the time that was left per move if the player were to follow the shortest W-path.

Appendix Table AT.8: Replication of Table 3, Controlling for Time Left per Move

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–30.105 (0.018)	–32.975 (0.150)	–54.813 (0.045)	–23.409 (0.065)	–9.002 (0.007)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	10.305	11.129	26.210	14.992	4.301
R^2	0.234	0.247	0.319	0.208	0.133
N	3,217,115,234	56,772,033	138,639,890	83,387,752	2,797,083,668

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.267 (0.005)	0.149 (0.025)	1.318 (0.031)	0.190 (0.053)	0.577 (0.012)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.452	0.188	1.352	0.588	0.498
R^2	0.277	0.279	0.306	0.344	0.271
N	259,240,947	5,014,246	32,452,348	3,829,648	164,138,685

Notes: See Table 3 in the main text. Note that conditional on the included fixed effects there is no independent variation in the time that was left per move.

Appendix Table AT.9: Replication of Table 4, Controlling for Time Left per Move

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–0.864 (0.048)		–0.218 (0.063)	
Other Title (÷ 100)		–0.818 (0.052)		–0.193 (0.065)
Grandmaster (÷ 100)		–1.188 (0.131)		–0.395 (0.199)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	$<$ 0.001	$<$ 0.001	$<$ 0.001	0.002
H_0: Grandmasters = Other Titled Players		0.008		0.332
Fixed Effects: Number of W- × D- × L-Moves × Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	6.039	6.039	2.748	2.748
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.367	0.367	0.412	0.412
N	212,295,223	212,295,223	25,771,678	25,771,678

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that the results above also control for the time that was left per move if the player were to follow the shortest W-path.

Appendix Table AT.10: Replication of Table 5, Controlling for Time Left per Move

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of Mistake						
Number of W-Moves (÷ 100)	–0.280 (0.001)			–1.245 (0.002)	–0.989 (0.003)		
Number of D-Moves (÷ 100)		0.812 (0.002)				0.844 (0.003)	
Number of L-Moves (÷ 100)			0.247 (0.003)				
Total Number of Moves (÷ 100)							–0.017 (0.001)
Fixed Effects:							
Number of D- × L-Moves	Yes	No	No	No	No	No	No
Number of W- × L-Moves	No	Yes	No	No	No	No	No
Number of W- × D-Moves	No	No	Yes	No	No	No	No
Total Moves × Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves × Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves × Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- × D- × L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.119	0.144	0.154	0.121	0.143	0.151	0.155
N	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223	212,295,223

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above also control for the time that was left per move if the player were to follow the shortest W-path.

Appendix Table AT.11: Replication of Table 2, Games with Long Time Controls Only

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of Choosing Move					
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM × W-Move (÷ 100)	–1.201 (0.014)		–1.086 (0.012)	–0.498 (0.018)		–0.428 (0.017)
DTM × L-Move (÷ 100)		0.017 (0.004)	0.021 (0.006)		0.027 (0.007)	0.033 (0.008)
W-Move (÷ 100)			74.326 (0.450)			54.858 (0.595)
Fixed Effects:						
Number of W- × D- × L-Moves × Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	10.161	0.360	9.093	11.728	0.579	8.425
R^2	0.426	0.346	0.426	0.569	0.362	0.569
N	93,544,167	8,837,271	102,381,438	10,269,201	3,052,879	13,322,080

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that the results above restrict attention to games played under classical and correspondence time controls.

Appendix Table AT.12: Replication of Table 3, Games with Long Time Controls Only

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–32.921 (0.116)	–32.921 (0.116)	–56.283 (0.238)	–26.030 (0.352)	–9.618 (0.045)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	11.129	11.129	27.167	15.847	4.398
R^2	0.246	0.246	0.320	0.214	0.147
N	92,878,586	92,878,586	4,281,066	3,048,972	80,503,389

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.148 (0.019)	0.148 (0.019)	0.794 (0.140)	0.194 (0.098)	0.339 (0.035)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.188	0.188	0.723	0.247	0.208
R^2	0.276	0.276	0.319	0.332	0.284
N	8,300,536	8,300,536	910,703	117,945	5,322,638

Notes: See Table 3 in the main text. The only difference between this table and that in the text is that the results above restrict attention to games played under classical and correspondence time controls.

Appendix Table AT.13: Replication of Table 4, Games with Long Time Controls Only

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–2.870 (0.535)		–1.309 (0.939)	
Other Title (÷ 100)		–2.570 (0.493)		–1.475 (1.055)
Grandmaster (÷ 100)		–11.482 (3.236)		0.000 (0.000)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	< 0.001	< 0.001	0.163	0.162
H_0: Grandmasters = Other Titled Players		0.006		0.162
Fixed Effects: Number of W- × D- × L-Moves × Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	4.545	4.545	1.212	1.212
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.418	0.418	0.674	0.674
N	6,563,054	6,563,054	739,633	739,633

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that the results above restrict attention to games played under classical and correspondence time controls.

Appendix Table AT.14: Replication of Table 5, Games with Long Time Controls Only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of Mistake						
Number of W-Moves ($\div$ 100)	–0.242			–0.921	–0.679		
	(0.003)			(0.009)	(0.015)		
Number of D-Moves ($\div$ 100)		0.558				0.704	
		(0.009)				(0.016)	
Number of L-Moves ($\div$ 100)			0.123				
			(0.015)				
Total Number of Moves ($\div$ 100)							–0.056
							(0.005)
Fixed Effects:							
Number of D- $\times$ L-Moves	Yes	No	No	No	No	No	No
Number of W- $\times$ L-Moves	No	Yes	No	No	No	No	No
Number of W- $\times$ D-Moves	No	No	Yes	No	No	No	No
Total Moves $\times$ Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves $\times$ Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves $\times$ Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- $\times$ D- $\times$ L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.129	0.151	0.163	0.128	0.151	0.159	0.172
N	6,563,054	6,563,054	6,563,054	6,563,054	6,563,054	6,563,054	6,563,054

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above restrict attention to games played under classical and correspondence time controls.

Appendix Table AT.15: Replication of Table 2, Board Positions with High Minimal DTM Only

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of Choosing Move					
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM $\times$ W-Move ($\div$ 100)	–0.450		–0.357	–0.474		–0.318
	(0.004)		(0.003)	(0.009)		(0.008)
DTM $\times$ L-Move ($\div$ 100)		0.034	0.141		0.053	0.168
		(0.003)	(0.004)		(0.007)	(0.009)
W-Move ($\div$ 100)			55.228			56.068
			(0.281)			(0.634)
Fixed Effects:						
Number of W- $\times$ D- $\times$ L-Moves $\times$ Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	15.365	0.795	11.064	13.552	1.295	9.381
R^2	0.379	0.330	0.394	0.435	0.387	0.443
N	92,066,019	39,282,812	131,348,831	22,051,140	9,811,399	31,862,539

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that the results above restrict attention board configurations in which the DTM of the shortest W-path is at least 50.

Appendix Table AT.16: Replication of Table 3, Board Positions with High Minimal DTM Only

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–23.424 (0.064)	–26.030 (0.352)	–31.354 (0.183)	–23.424 (0.064)	–9.985 (0.046)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	14.992	15.847	28.223	14.992	8.495
R^2	0.208	0.214	0.211	0.208	0.177
N	89,336,853	3,048,972	5,832,993	89,336,853	72,891,541

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.166 (0.013)	0.084 (0.062)	1.626 (0.119)	0.182 (0.142)	0.423 (0.027)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	≤ 10 Total Moves	DTM ≥ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.504	0.246	2.059	0.914	0.496
R^2	0.278	0.295	0.329	0.337	0.272
N	37,575,901	1,282,479	2,686,109	220,360	24,976,281

Notes: See Table 3 in the main text. The only difference between this table and that in the text is that the results above restrict attention board configurations in which the DTM of the shortest W-path is at least 50.

Appendix Table AT.17: Replication of Table 4, Board Positions with High Minimal DTM Only

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–2.228 (0.331)		–0.032 (0.239)	
Other Title (÷ 100)		–2.052 (0.367)		–0.272 (0.177)
Grandmaster (÷ 100)		–3.342 (0.663)		1.436 (1.271)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	< 0.001	< 0.001	0.894	0.157
H_0: Grandmasters = Other Titled Players		0.089		0.182
Fixed Effects: Number of W- × D- × L-Moves × Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	20.525	20.525	3.593	3.593
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.470	0.470	0.487	0.487
N	12,927,786	12,927,786	2,121,748	2,121,748

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that the results above restrict attention board configurations in which the DTM of the shortest W-path is at least 50.

Appendix Table AT.18: Replication of Table 5, Board Positions with High Minimal DTM Only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Probability of Mistake						
Number of W-Moves (÷ 100)	–1.465			–2.986	–1.931		
	(0.005)			(0.006)	(0.010)		
Number of D-Moves (÷ 100)		1.524				1.338	
		(0.006)				(0.010)	
Number of L-Moves (÷ 100)			0.235				
			(0.009)				
Total Number of Moves (÷ 100)							–0.317
							(0.006)
Fixed Effects:							
Number of D- × L-Moves	Yes	No	No	No	No	No	No
Number of W- × L-Moves	No	Yes	No	No	No	No	No
Number of W- × D-Moves	No	No	Yes	No	No	No	No
Total Moves × Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves × Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves × Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- × D- × L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.224	0.237	0.247	0.224	0.237	0.240	0.252
N	12,927,786	12,927,786	12,927,786	12,927,786	12,927,786	12,927,786	12,927,786

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above restrict attention board configurations in which the DTM of the shortest W-path is at least 50.

Appendix Table AT.19: Replication of Table 2, Weighting All Observations Equally

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of Choosing Move					
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM × W-Move (÷ 100)	–1.171		–0.931	–0.696		–0.315
	(0.001)		(0.001)	(0.002)		(0.001)
DTM × L-Move (÷ 100)		0.024	–0.069		0.066	0.002
		(0.000)	(0.000)		(0.001)	(0.001)
W-Move (÷ 100)			51.783			28.183
			(0.035)			(0.044)
Fixed Effects:						
Number of W- × D- × L-Moves × Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	6.186	0.457	5.734	6.743	0.631	5.404
R^2	0.278	0.183	0.274	0.348	0.175	0.340
N	3,457,878,398	296,522,573	3,754,400,971	398,856,135	111,905,262	510,761,397

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that the results above are based on equally weighted observations.

Appendix Table AT.20: Replication of Table 3, Weighting All Observations Equally

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–19.846 (0.012)	–22.446 (0.044)	–54.658 (0.020)	–20.392 (0.022)	–7.728 (0.003)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	5.723	6.178	21.212	9.854	3.009
R^2	0.112	0.127	0.293	0.128	0.071
N	3,435,257,516	92,878,586	148,599,747	89,336,853	2,986,665,574

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.380 (0.002)	0.211 (0.005)	1.499 (0.008)	0.240 (0.012)	0.689 (0.003)
Fixed Effects: Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.456	0.174	1.248	0.487	0.502
R^2	0.189	0.188	0.245	0.255	0.202
N	277,507,532	8,300,536	34,762,492	4,095,758	175,675,645

Notes: See Table 3 in the main text. The only difference between this table and that in the text is that the results above are based on equally weighted observations.

Appendix Table AT.21: Replication of Table 4, Weighting All Observations Equally

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Titled Player (÷ 100)	–0.367 (0.052)		–0.005 (0.048)	
Other Title (÷ 100)		–0.342 (0.055)		0.006 (0.050)
Grandmaster (÷ 100)		–0.589 (0.137)		–0.095 (0.144)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	$<$ 0.001	$<$ 0.001	0.923	0.800
H_0: Grandmasters = Other Titled Players		0.092		0.508
Fixed Effects: Number of W- × D- × L-Moves × Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	5.751	5.751	2.559	2.559
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.301	0.301	0.271	0.271
N	226,955,095	226,955,095	27,600,514	27,600,514

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that the results above are based on equally weighted observations.

Appendix Table AT.22: Replication of Table 5, Weighting All Observations Equally

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Probability of Mistake			
Number of W-Moves ($\div$ 100)	–0.271			–1.223	–0.985		
	(0.000)			(0.001)	(0.002)		
Number of D-Moves ($\div$ 100)		0.800				0.814	
		(0.001)				(0.001)	
Number of L-Moves ($\div$ 100)			0.254				
			(0.001)				
Total Number of Moves ($\div$ 100)							–0.016
							(0.000)
Fixed Effects:							
Number of D- $\times$ L-Moves	Yes	No	No	No	No	No	No
Number of W- $\times$ L-Moves	No	Yes	No	No	No	No	No
Number of W- $\times$ D-Moves	No	No	Yes	No	No	No	No
Total Moves $\times$ Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves $\times$ Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves $\times$ Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- $\times$ D- $\times$ L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.105	0.130	0.140	0.106	0.129	0.138	0.142
N	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095	226,955,095

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that the results above are based on equally weighted observations.

Appendix Table AT.23: Replication of Table 2, TWIC Data

	(1)	(2)	(3)	(4)	(5)	(6)
			Probability of Choosing Move			
	W-Moves	L-Moves	W- & L-Moves	W-Moves	L-Moves	W- & L-Moves
DTM $\times$ W-Move ($\div$ 100)	–1.443		–1.216	–1.420		–1.012
	(0.030)		(0.025)	(0.145)		(0.111)
DTM $\times$ L-Move ($\div$ 100)		0.002	0.047		0.000	0.036
		(0.006)	(0.030)		(0.001)	(0.050)
W-Move ($\div$ 100)			88.962			80.414
			(1.525)			(5.279)
Fixed Effects:						
Number of W- $\times$ D- $\times$ L-Moves $\times$ Complexity of Other Moves	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes
Board Configurations	All	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$	$\mid D \mid = 0$
Mean of LHS Variable (%)	14.779	0.168	11.887	13.190	0.316	9.032
R^2	0.533	0.369	0.545	0.508	0.347	0.573
N	5,132,343	1,177,248	6,309,591	853,980	296,615	1,150,595

Notes: See Table 2 in the main text. The only difference between this table and that in the text is that results above are based on data from The Week in Chess.

Appendix Table AT.24: Replication of Table 3, TWIC Data

Panel A: W-Moves

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
Percentile Rank (÷ 100)	–46.902	–47.100	–66.285	–38.238	–11.261
	(0.212)	(0.217)	(0.499)	(0.506)	(0.114)
Fixed Effects:					
Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	15.603	15.661	28.323	17.258	4.562
R^2	0.321	0.322	0.392	0.269	0.197
N	5,064,773	4,583,088	339,083	958,023	4,311,578

Panel B: L-Moves

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
Percentile Rank (÷ 100)	0.076	0.078	0.491	–0.070	0.156
	(0.028)	(0.029)	(0.151)	(0.094)	(0.058)
Fixed Effects:					
Player × Choice Set	Yes	Yes	Yes	Yes	Yes
Sample	Full	Long Time Controls	$\leq$ 10 Total Moves	DTM $\geq$ 50	Excl. Simplest Move
Mean of LHS Variable (%)	0.083	0.083	0.329	0.029	0.093
R^2	0.248	0.243	0.287	0.486	0.294
N	1,119,866	1,023,218	76,968	12,417	733,321

Notes: See Table 3 in the main text. The only difference between this table and that in the text is that results above are based on data from The Week in Chess.

Appendix Table AT.25: Replication of Table 4, TWIC Data

	(1)	(2)	(3)	(4)
	Probability of Mistake			
Top 25% of Players (÷ 100)	–3.260		–0.032	
	(0.336)		(0.032)	
75th to 99th Percentile of Players (÷ 100)		–3.184		–0.031
		(0.342)		(0.033)
Top 1% of Players (÷ 100)		–4.873		–0.051
		(0.605)		(0.046)
Hypothesis Tests (p-value):				
H_0: No Differences between Players	$<$ 0.001	$<$ 0.001	0.313	0.427
75th to 99th Percentile = Top 1%		0.005		0.683
Fixed Effects:				
Number of W- × D- × L-Moves				
× Complexity of Moves	Yes	Yes	Yes	Yes
Mean of LHS Variable (%)	5.133	5.133	0.786	0.786
Board Configurations	All	All	$\mid D \mid = 0$	$\mid D \mid = 0$
R^2	0.520	0.520	0.650	0.650
N	499,331	499,331	65,113	65,113

Notes: See Table 4 in the main text. The only difference between this table and that in the text is that results above are based on data from The Week in Chess. Since the TWIC data do not consistently list players' titles, we rely on ELO ratings to differentiate players by skill.

Appendix Table AT.26: Replication of Table 5, TWIC Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Probability of Mistake			
Number of W-Moves (÷ 100)	–0.140			–0.590	–0.269		
	(0.007)			(0.014)	(0.022)		
Number of D-Moves (÷ 100)		0.424				0.432	
		(0.015)				(0.025)	
Number of L-Moves (÷ 100)			0.073				
			(0.023)				
Total Number of Moves (÷ 100)							0.019
							(0.010)
Fixed Effects:							
Number of D- × L-Moves	Yes	No	No	No	No	No	No
Number of W- × L-Moves	No	Yes	No	No	No	No	No
Number of W- × D-Moves	No	No	Yes	No	No	No	No
Total Moves × Number of L-Moves	No	No	No	Yes	No	No	No
Total Moves × Number of D-Moves	No	No	No	No	Yes	No	No
Total Moves × Number of W-Moves	No	No	No	No	No	Yes	No
Fraction of W- × D- × L-Moves	No	No	No	No	No	No	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minimal DTM among W-Moves	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.339	0.339	0.343	0.337	0.341	0.343	0.368
N	536,674	536,674	536,674	536,674	536,674	536,674	536,674

Notes: See Table 5 in the main text. The only difference between this table and that in the text is that results above are based on data from The Week in Chess.