

How Men and Women Respond to Failure: Evidence from Chess Tournaments

Julia Francesca Engel *

November 3, 2025

- Job Market Paper -

[\[Please click here for the latest version.\]](#)

Abstract

Setbacks are inevitable in professional life, yet we know little about how people respond to them. I study how men and women respond to failure in a highly competitive setting, using 286,140 games from 3,523 titled chess players in a weekly tournament series. The weekly, 11-round structure and fixed player pool allow me to track both short- and long-term responses to losses. After a loss, men play less accurately, choose riskier openings, and are more likely to quit the tournament. In contrast, women maintain or improve their performance, showing greater resilience. Over the long run, however, men and women are equally likely to reenter tournaments and invest in skill development, conditional on past outcomes. These findings suggest that short-term emotional responses differ by gender, but longer-term competitive choices are primarily skill- and success-driven.

Keywords: Competition, Gender Differences, Willingness to Compete, Chess

JEL Codes: J16, C93

*Kiel University. Acknowledgements: My deepest gratitude to Patrik Gränsmark and Christer Gerder for sharing their opening classification dataset with me. My sincerest thanks goes to Kiel University's Institute for Statistics and Econometrics for providing the computing power necessary to take on this project. I would like to thank Pau Brinkhaus Tort, Maan Yazgi, Torben von Wehgeln, and Lasse Clause for excellent research support. I would like to thank Jens Ruhose, Maria Cubel, Patrick Nüß, Santiago Sánchez-Páges, José De Sousa, Maya Jalloul, and Loren Fryxell for their comments and valuable feedback.

1 Introduction

Failure is a normal part of any career: job applicants get turned down and promotions go to others. Advancing to senior managerial levels requires not just competing once, but returning to competition again and again. This is particularly relevant in light of the persistent gender gap in top positions, where men continue to outnumber women (Bertrand and Hallock, 2001). A large literature shows that women are less willing to enter competitive environments than men (Gneezy et al., 2003; Niederle and Vesterlund, 2007; Wasserman, 2023), yet we know much less about how men and women respond to failure once they are already competing. Understanding these reactions is crucial, since discouragement at the top level can limit women’s representation in leadership and affect opportunities for other women in the organization (Flabbi et al., 2019). Laboratory experiments are difficult to generalize to the dynamics of competition among those pursuing top managerial roles, as participants are often students who have not experienced high-stakes selection processes (Gill and Prowse, 2014; Alnamlah and Gravert, 2020). This limitation motivates the need for real-world data that captures competitive behavior in actual high-stakes career contexts.

Studying responses to failure with labor market data is challenging. Measuring the impact of failure on job performance is difficult because performance is rarely standardized across firms or roles (Ahmed et al., 2013; Schröder-Hansen and Hansen, 2023), evaluations may differ by gender (Doldor et al., 2019), and constructing counterfactuals is complicated by institutional differences. Another suitable option could be job application data where failure is well-defined (Bapna et al., 2021; Brands and Fernandez-Mateo, 2017), but measuring subsequent job performance is again made difficult across differing jobs. In short, existing data sources rarely combine three features necessary to study reactions to failure: (i) a clearly defined outcome of success or failure, (ii) a standardized performance measure, and (iii) the ability to observe whether individuals re-enter competition.

To overcome these limitations, I use chess tournament data from the top 1% of chess players in the world to study how disappointing outcomes affect subsequent game performance and tournament reentry. This setting offers several advantages: matches have clear outcomes that allow precise identification of failure, Elo ratings provide an objective measure of skill, and individuals can be tracked over time to observe re-entry decisions. I analyze two years of weekly data from Chess.com’s Titled Tuesday series, covering 286,140 games played by 3,508 chess title-holding players. From these data, I reconstruct each player’s individual tournament trajectory and placements across all events. Furthermore, modern chess computers allow me to measure the quality of every move. I analyze over 20 million moves and compute the advantage lost between a player’s move and the engine’s optimal move as an objective measure of cognitive performance. Additionally, I classify opening strategies into risk profiles using expert assessments, offering a second behavioral margin beyond move accuracy. Because tournaments consist of multiple rounds and occur weekly, I can study both short-run responses to losses *within* tournaments and longer-run re-entry choices *between* tournaments. With this data, I follow players not only through a single tournament, but across their entire participation and chess activity on the platform in the entire two-year period.

I document two main findings. First, in the short run, men’s performance deteriorates after a loss: their move accuracy decreases and they adopt riskier opening strategies. By contrast, women’s move accuracy improves following a loss, suggesting greater performance stability in the face of setbacks. Second, in the long run, men and women are equally likely to re-enter tournaments once past performance is taken into account. Successful players of both genders re-enter sooner and invest more in practice, while poorly performing players wait longer. This behavior is consistent

with updating based on objective feedback. Overall, the results show that failure has immediate but gender-differentiated effects on performance, while influencing men’s and women’s competitive choices in similar ways over the longer run. The objective nature of chess feedback ensures that players’ performance is accurately represented. Under these conditions, men and women who initially competed do not differ in their long-run willingness to compete again, contrasting with the narrative that women are systematically less willing to remain in competition. These findings suggest that objective feedback environments can help narrow the gender gap and ensure that strong performers continue to opt into competition at the expert level.

This paper contributes to the literature on gender and competition by focusing on how individuals respond to failure. While prior laboratory experiments have shown that failure can reduce subsequent competitive effort, these studies typically rely on student participants and may not fully capture competitive behavior in professional settings (Gill and Prowse, 2012; Buser, 2016; Alnamlah and Gravert, 2020). Previous work also largely measures outcomes between stages of a task, limiting insight into short-run performance fluctuations. Similarly, most field studies can observe whether individuals reenter competitions or who ultimately “wins,” but they cannot measure performance quality or track short-run progress within the competition. Job application data illustrate this limitation: candidates receive varying feedback depending on how far they advance, but such nuances are rarely observable. My paper leverages a unique field setting with expert chess players, allowing me to examine both the immediate effects of failure on cognitive performance and the longer-run impact on reentry decisions and skill investment. A study of two years of reentry choices shows that these decisions are informed by feedback from previous competitive efforts, in line with the broader literature on feedback provision and shrinking gender gaps in competition.

The paper is organized as follows: Section 2 reviews the literature on gender and competition. Section 3 describes the institutional setting. Section 4 presents the data and variables. Section 5 outlines the identification strategy. Section 6 reports the results on short-run and long-run outcomes, which are then discussed in Section 7. Section 8 concludes.

2 Behavior in Competitive Settings

Research on gender differences in competitive behavior primarily stems from laboratory experiments. These studies typically compare behavior across stages or treatments in static settings, with time dynamics restricted to short-run outcomes. Few studies capture repeated competitive behavior over longer periods, particularly in professional contexts. In contrast, natural field settings offer the potential to observe long-run behavior and richer performance measures, providing a complementary perspective to lab experiments.

Laboratory experiments consistently find that introducing competition boosts performance for both men and women (Gneezy et al., 2003; Gneezy and Rustichini, 2004). At the same time, women tend to opt out of competing at higher rates than men, especially when the task is perceived as male-favoring (Markowsky and Beblo, 2022; Niederle and Vesterlund, 2007; Shurchkov, 2012). Even when a woman is of high ability and has the opportunity to be successful, she may abstain from competing. Institutional changes, such as allowing a familiarity with the setting (Buser, 2016; Cotton et al., 2013; Niederle and Yestrumskas, 2008), affirmative action measures (Niederle et al., 2013) may help narrow the gap and improve self-selection into competition. Lab experiments with performance feedback to participants suggest it may be an avenue to narrow the gender gap in willingness to compete (Berlin and Dargnies, 2016; Ertac and Szetes, 2011; Piasenti et al., 2023;

Shastri et al., 2020; Wozniak et al., 2016). I extend this line of work by linking immediate performance outcomes to subsequent participation decisions, providing evidence on whether competitive experiences shape the gender gap over time.

The paper further contributes to the broader literature on the impact of success and failure on performance. Gill and Prowse (2012) found that individuals decrease their effort in a sequential tournament in response to an opponent's strong effort. Later work suggests that losses have a negative effect on both men's and women's performance (Gill and Prowse, 2014). Conversely, early victories have been found to boost later performances (González-Díaz and Palacios-Huerta, 2016; Koizumi, 2024). My work builds on these findings in the field with professionals. Using chess data allows me to study these behavioral outcomes at a greater detail than previous lab work has. I have detailed data on the quality of work (chess play) being performed against a neutral benchmark and can show that failure not only impacts performance, but also players' willingness to take risks.

In general, my findings align with the literature that failure and rejection decrease willingness to compete again. With an RDD design and Dutch Math Olympiad data Buser and Yuan (2019) find girls who narrowly missed the cut-off mark in knock-out rounds are less likely to reenter to competition than their male peers in the following year. Similar outcomes are reported in the sports literature (Haenni, 2019; Wozniak, 2012). Brands and Fernandez-Mateo (2017) find that women are less likely to reapply to firms that previously rejected them. For female politicians, Wasserman (2023) finds that women are less likely than men to run again but that this gender gap declines with experience in politics. My contribution is to offer a more detailed perspective by analyzing weekly reentry decisions over two years, in a setting where feedback is both immediate and unambiguous, ensuring that players' responses to relative placement can be cleanly identified.

Chess offers a promising quasi-experimental setting for studying competitive behavior (Backus et al., 2023; De Sousa and Niederle, 2022; De Sousa and Hollard, 2023). Its key features mirror those of high-level management, which is characterized as a competitive, cognitively demanding, and male-dominated environment (Autor and Handel, 2013; Bertrand and Hallock, 2001; Goldin, 2014; Niessen-Ruenzi and Ruenzi, 2019). As in the labor market for high-level management, chess features few expert female players suggesting a leaky pipeline (De Sousa and Niederle, 2022). Chess is uniquely suited to explore questions surrounding gender and competition as it is one of the few competitive settings with mixed-gender competition (Palacios-Huerta, forthcoming). A further advantage of chess is that game outcomes are the result of skill and cognitive effort. The analysis of chess games has received attention in the labor literature as professional play features a numerical skill measure, randomness in opponent allocation and clean data on games and players. Taken together, these features make chess an ideal field setting for studying gender differences in behavior.

The analysis of chess games suggests that women perform worse against men than similarly skilled women (Backus et al., 2023; Maass et al., 2008; Smerdon et al., 2020). There are further gender differences in playing styles and risk attitudes: Men's play style at the beginning of a match is more aggressive than that of women (Gerdes and Gränsmark, 2010; Dreber et al., 2013a). However, shorter time formats close the gender gap in risk taking (Dilmaghani, 2021). Differences are also observable in the playing times of lost games, men are more reluctant to resign when playing against women compared to men (Dilmaghani, 2022; Backus et al., 2023). Further research has centered around the cognitive dimension and playing quality of chess games. Künn et al. (2023) can quantify how much pollution impairs concentration by exploiting variation in air quality of a chess tournament. To study the effects of remote work on cognitive performance Künn et al. (2022) compare within-player differences in playing quality for a high-stakes online tournament

versus in-person tournament. [Strittmatter et al. \(2020\)](#) use chess to document the evolution of cognitive over the life cycle. Research by [Cubel et al. \(work-in-progress\)](#) quantifies how exposure to conflict negatively affects the cognitive performance of civilians.

Taken together, the existing literature highlights both the potential and the limitations of current approaches to studying gender differences in competition. Laboratory experiments have documented systematic patterns in performance, risk-taking, and willingness to compete, while field studies emphasize the influence of context, feedback, and persistence over time. This study contributes to this body of work by examining a professional field setting in which outcomes and performance are precisely measured and competition is mixed-gender. Importantly, the chess setting offers unambiguous feedback on outcomes and performance, reducing uncertainty about whether players actually receive and internalize information about their successes and failures. The setting allows for an investigation of how failure shapes subsequent performance and informs reentry decisions, offering complementary evidence to laboratory experiments on the dynamics of gender differences in competition in high-stakes environments.

3 Institutional Background

This paper introduces chess tournament data from a cash-prize tournament series of the world’s largest online chess playing platform. The Chess.com Titled Tuesday event has been held weekly since April 7, 2020, with participation restricted to players holding an over-the-board chess title, ensuring a sample of the world’s top players who meet well-documented skill requirements ([FIDE, 2021](#)). The dataset begins when the Titled Tuesday tournament became a weekly series. The observation period ends in April 2022 following a ban of the platform by Russian internet providers in response to chess.com’s public criticism of the Russian invasion in Ukraine.¹

Each tournament consists of 11 rounds and takes approximately 3 hours from beginning to conclusion.² Due to the limited time frame, players play a speed chess variant. Players have 3 minutes each to make their moves and receive an additional second for every move, a popular format for online play. Unlike classical games, which allow 90 minutes for 40 moves, this format requires balancing rapid calculation, time management, and composure under pressure ([Künn et al., 2023](#)), skills related to but distinct from pure chess ability. The format emphasizes performance under time constraints.

The winner of each tournament receives 750 USD, with additional lower cash prizes for places 2-4 as well as the best woman. The tournament uses a point system to determine the winner. Matches award 1 point for a win, 0.5 for a draw, and 0 for a loss. The player with the highest total points after the final round wins. Tie-breaking rules apply in case of a point tie (see [Appendix A.1.2](#)).

Player pairings are determined by the “Swiss System”. Players are quasi-randomly paired in round 1 and paired by tournament standings (sum of points achieved) in all subsequent rounds.³ In subsequent rounds winners play winners and losers play losers, with random pairings within these

¹One of the aims is to study reentry patterns and a subset of players faced additional barriers after April 2022. While the ban can be circumvented, Russian players are one of the top five nationalities in the sample and their (involuntary) exclusion may change the tournament composition and affect their own and others’ reentry choices. See [Appendix A.1.1](#) for details.

²Between April 7, 2020 and October 20, 2020, the Titled Tuesday tournament was a 10-round tournament and was then expanded to 11 rounds.

³Players are seeded by their ranking relative to all other entrants. They cannot influence this process. The competitor pool is ranked and divided into two halves. The top players of both halves are paired, the second best players of each half and so forth until all are paired.

groups. Starting in round 3, players’ tournament trajectories diverge, but allocation within the pool remains random, and players cannot influence their opponents (Figure 1).

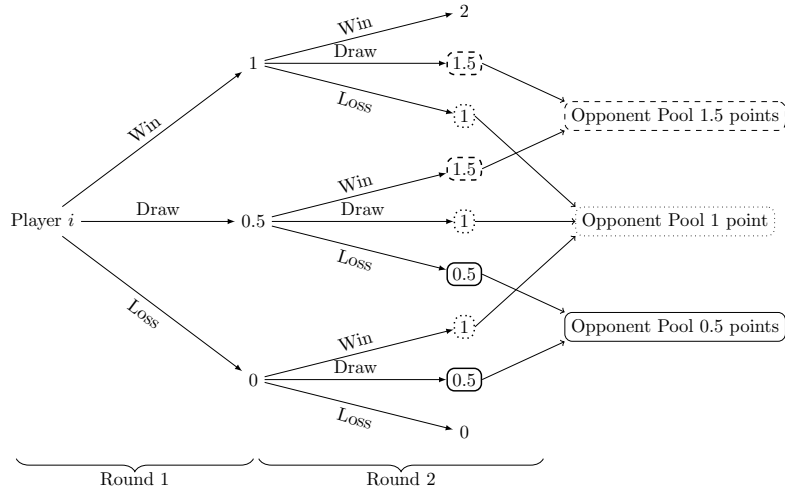


Figure 1: Swiss System Opponent Pairings

Note: The figure illustrates pairings in the first two rounds of a Swiss-system tournament. Pairings are determined by the points each player has accumulated so far. After the second round, players with differing performance trajectories may be placed in the same opponent pool.

In competitive chess, players are ranked by the Elo system (Elo, 1978). This system, or slight variations thereof, is commonly used in sports to rank players (e.g. the Universal Tennis Rating or esports). Ratings reflect skill relative to other rated players, adjusting upward for wins or draws against stronger opponents and downward for losses or draws against weaker opponents. Ratings are central to a chess player’s identity (González-Díaz et al., 2024) and are displayed in every match, capturing initial skill differences. All games in the Titled Tuesday tournament are rated to further incentivize players. Crucially for this setting, chess ratings are displayed in every match and can capture the initial skill difference between players (see Appendix A.1.3).

4 Data

This section introduces the dataset and discusses key dependent and independent variables. The primary outcomes of interest are reactions to failure, in particular how failure affects the quality of work being performed and risk preferences. The secondary outcomes center around tournament reentry behavior and skill investment over a longer time horizon.

4.1 Dataset

The sample consists of two years’ data from the chess.com Titled Tuesday tournament series from April 7, 2020 to April 24, 2022. In total, it covers 286,140 unique games played by 4,745 accounts across 114 separate tournament events. Because each tournament comprises multiple rounds, the data forms a panel of rounds (r) nested within tournaments (t) across the T total tournaments.. Chess.com requires a clear username policy for accounts to be eligible for cash prizes. Usernames are then linked to FIDE administrative records and supplemented with national federation, gender, and age information. Of the 4,745 unique accounts in the dataset, 3,523 accounts (74%) could be linked to existing FIDE profiles.⁴

⁴Since chess.com’s title verification occurs manually, the unidentified accounts possess expert-level chess knowledge.

The game data are structured so that each identified player in a given round of a tournament represents one observation; if both players are identified, this produces two observations per game. The tournament files, provided in standard portable game notation (PGN) format, include each match’s moves, round, outcome, piece color, usernames, and ratings. Players see each other’s ratings, titles, usernames, profile pictures, and current tournament standings during play (see Appendix A.1.4). I use the chess.com application programming interface (API) to download this additional information and reconstruct each player’s tournament trajectory, including indicators for losses in previous rounds. Summing points and calculating tiebreaks (Appendix A.1.2) recovers each player’s overall placement within the tournament.

The initial dataset comprises 286,140 games, yielding two observations per game (White and Black) for a total of 592,280 observations. For the analysis, I include only identified accounts linked to a FIDE profile with complete information on gender, birth year, and chess federation. Observations of identified players facing unidentified opponents remain, while observations from anonymous accounts are excluded. I drop accounts with fewer than 11 recorded games, as these represent incomplete tournaments, and omit games with fewer than 10 moves to accurately identify opening choices. Since the analysis focuses on the role of failure, I also exclude observations from the first three rounds to allow sufficient variation in tournament trajectories. The final sample contains 312,822 game-level observations.

4.2 Dependent Variables

4.2.1 Cognitive Performance

The average chess game in the dataset lasts 43 moves, which requires players to identify the best move, or at least a good move, each time. The best moves are found by forwards iteration and predicting the opponent’s response for several moves ahead. This process is cognitively demanding and requires players to solely focus on the game in front of them.

For each player in each game, I measure cognitive performance by comparing their moves to those suggested by a computer chess engine. Using Stockfish 14.1 at depth 20, I benchmark each move against a neutral, computationally optimal reference (Acher and Esnault, 2016; Backus et al., 2023; Künn et al., 2022, 2023; Smerdon, 2022; Strittmatter et al., 2020). Specifically, for each tournament t , round r , move number m , and player i , the chess engine calculates the relative advantage of White or Black (C_{irtm}) with the best possible move. The player’s actual move produces a realized advantage P_{irtm} . I define cognitive performance as the difference between the computer-optimal and player-achieved advantages. This difference, always nonnegative, is called the centipawn loss and quantifies the player’s lost advantage from not finding or playing the computer’s best move (Equation 1).

$$CPLoss_{irtm} = C_{irtm} - P_{irtm} \in [0; \infty] \quad (1)$$

A low centipawn loss indicates moves close to the computer’s best choice, while a high value signals significant mistakes. Accordingly, a game’s cognitive performance is higher when the centipawn loss is lower. For an example and additional detail on how this metric is composed, see Appendix A.2.1. I calculate the average centipawn loss per move for each game. For an individual i playing in round r of tournament t , this measure is comprised of the total sum of centipawn loss per move over a game, divided by the number of moves (M_{irt}).

$$\overline{CPLoss}_{irt} = \frac{\sum_{m=1}^{M_{irt}} C_{irtm} - P_{irtm}}{M_{irt}} \in [0; \infty] \quad (2)$$

The average centipawn loss is 70, which is higher than in previous papers (Backus et al., 2023; Smerdon, 2022; Künn et al., 2023). This is consistent with the speed chess format where players have only three minutes to play, and receive an additional second for each move. To avoid losing on time, players might play “second-best” moves instinctively rather than spending a minute to find the optimal move. The distribution of errors across moves aligns with prior findings (Künn et al., 2023): few mistakes occur in the opening, error magnitude rises in the middlegame, and peaks in long endgames under mounting time pressure (see Figure A.5).

For cognitive performance to be meaningful, accurate play should be associated with a increased winning probability. Figure 2 plots the (correlational) relationship of centipawn loss and its square on the probability of winning the match, additionally controlling for player and opponent strength. The relationship is clearly negative: an increase of the average centipawn loss by one is associated with a 0.5pp reduction winning probability.

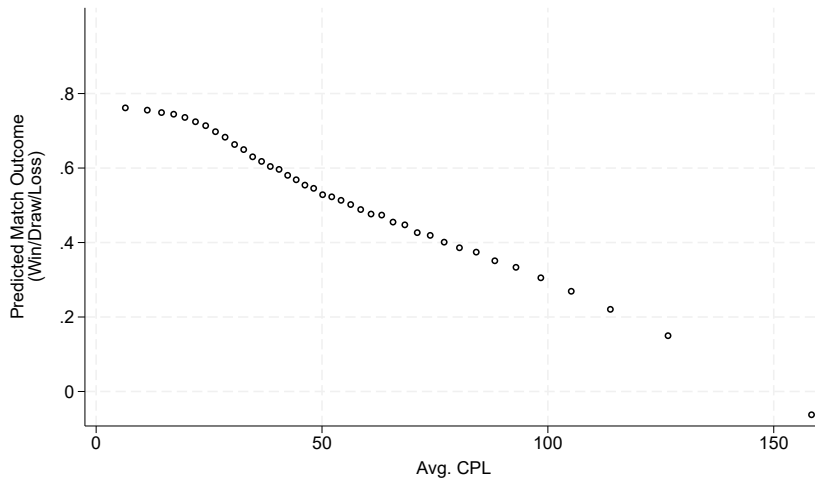


Figure 2: Predicted Game Outcomes on Average (Capped) Centipawn Loss

Note: The figure presents a binscatter in 40 bins of the predicted number of points won in a game by average centipawn loss for the full analytical sample. Every point represents the within-quantile mean. Predictions are based on a sample of games played by identified players, excluding the first three rounds. The score is regressed on the average centipawn loss and its square. Additional controls include for the piece color, the player’s own rating, the rating difference to their opponent, their age, the opponent’s title and profile picture content. The prediction is based on a regression with player, round and tournament fixed effects, the coefficients are reported in Table A.3.

4.2.2 Risk in Opening Choices

The first 15 moves of the game (the opening) determines the game’s direction. Titled and experienced chess players maintain an “opening repertoire”: a set of openings and responses they know well and feel comfortable playing as either White or Black. Because mastering specific openings requires substantial time and effort, players select openings that fit their personality, style of play, and risk preferences (Gerdes and Gränsmark, 2010; Dreber et al., 2013a). Openings are highly studied, with well-established best responses to common moves, and selecting them relies as much on memorized preparation as on pure chess skill.

In chess, opening styles are broadly classified as “aggressive” or “solid” (defensive). Aggressive

openings present early threats and often sacrifice smaller pieces such as pawns to create attacking opportunities and imbalance. Solid openings prioritize positional integrity and piece safety, minimizing immediate risk while limiting early attacks. The success of a style depends on player skill and opponent response, and the varying levels of risk make high-level players well-suited for analyzing risk preferences (Dreber et al., 2013a; Gerdes and Gränsmark, 2010). White, as the first mover, can dictate the initial flow of the game, while Black responds strategically to these choices.

Openings are classified in the Encyclopedia of Chess Openings (ECO), which assigns 499 distinct codes based on the first several moves of a game. Each ECO code represents a specific sequence of moves up to the 14th move and is rated by experts as “aggressive”, “solid”, or “ambiguous” (Gerdes and Gränsmark, 2010). Approximately two-thirds of ECO codes have a clearly identifiable style. Using the tournament game files, I classify each opening by its ECO code and merge it with the expert database to create a binary variable indicating whether an opening is deemed “aggressive”. Because players choose from a known repertoire of openings and these choices are observable for each game, I can track how risk preferences vary within and across tournaments, controlling for player skill and other confounders. Prior work finds gender differences in opening choices: women tend to select less risky openings than men, and men are more likely to play aggressively against female opponents.

4.2.3 Quitting the Tournament

Because the PGNs contain complete tournament records, I can identify players who quit a tournament entirely or skipped a round. For a given round r that is not the final round, a subsequent record in round $r + 1$ should exist; missing a record signals that the player quit. In chess, abandoning a tournament is generally considered poor sportsmanship, but in this online tournament without a registration fee, quitting is as simple as closing a browser window. Across 428,848 eligible observations, only 23,588 games (5.5%) involved a player quitting, highlighting both the low frequency of such behavior and the seriousness of participants in this elite sample.

4.2.4 Time to Reentry

The weekly schedule of the Titled Tuesday tournament and the well-defined player pool make it well-suited to study reentry behavior over time. For each player i , I record their first appearance in the tournament series after it became weekly. I then calculate the number of days until their next appearance to measure reentry timing. This gap is calculated for all of a player’s tournament appearances.

4.2.5 Platform Activity

Chess.com maintains an archive of each player’s activity, including a complete record of their games on the platform. Using the chess.com API, I reconstruct each player’s entire account activity between April 7, 2020 and April 20, 2022. In particular, I collect data on the number of bullet and blitz games, the platform’s most popular speed chess formats. I aggregate these on a weekly basis, creating two variables that capture a player’s (online) engagement with speed chess between two competition opportunities.

4.3 Covariates

4.3.1 Elo Rating

The tournament data files contain each player’s Chess.com Elo rating at the start of their matches. As described in Section 3, Elo ratings rank players by skill, assuming chess ability broadly follows a

bell-curve distribution (Elo, 1978). Ratings on the Chess.com range from 0 to approximately 3350, with new accounts starting at 800, 1200, or 1800 depending on claimed skill. Ratings quantify both a player’s own skill and their gap to opponents, confirming the expert nature of the sample: an account above 2000 points is in the top percentile of chess.com users, and the average account in the sample is 2506 (Table 1). Elo ratings are strongly predictive of tournament outcomes and move accuracy.⁵

Elo ratings can also predict the probability of a player winning a match based on the rating difference (see Appendix A.1.3). Players with equal ratings are expected to have a 50% win probability each, while a 300-point advantage increases the higher-rated player’s expected wins to 85%. The final dataset includes both each player’s rating and the rating difference to their opponent for every match. Using this difference, I identify the presumed favorite of each game.

4.3.2 Opponent Characteristics

During matches, players have access to several pieces of information in addition to the board: the opponent’s username, rating, chess title, and profile picture. Hovering over the username reveals a real name if provided. Information on an account’s titles and profile pictures is downloaded via the chess.com API. Profile pictures were manually coded into four categories: “male” if a male-presenting individual is depicted, “female” if a female-presenting individual is depicted, “other” for multiple people, icons, or other imagery, and “no picture” if none is provided. Pictures could have changed over time between 2020 and 2022. While historical account data is unavailable, scraping profile pictures between November 2022 and June 2024 shows players changed pictures about 18 times over this period. Assuming similar behavior in the earlier period, picture changes are unlikely to materially affect the analysis. Ratings and chess titles are more informative of skill than profile pictures.

Profile pictures may convey an opponent’s gender, potentially influencing players’ behavior (Gerdes and Gränsmark, 2010; Dreber et al., 2013a; Backus et al., 2023; Smerdon et al., 2020). However, 80% of female players in the sample hold a female-specific title (indicated by a “W” prefix, see Appendix A.1.5), which is displayed next to the username and does not change over time. This provides a stable, unambiguous signal for gender, allowing identification of behavioral responses to female opponents independent of profile picture changes.

4.3.3 Performance Metrics

A player’s overall tournament placement is determined by the raw sum of points after the final round together with the tiebreak points. When making re-entry decisions, players likely consider their past tournament experience as a whole instead of focusing on individual rounds. To capture this, this paper uses a variety of individual and relative measures to capture tournament-level performance.

On an individual level, two commonly used measures assess chess performance (FIDE, 2021). The first is the percentage score, defined as the sum of points scored across all rounds of tournament t divided by the number of rounds, as in Equation 3. It ranges from 0 to 1, with 1 indicating all wins, 0.5 reflecting as many draws as wins/losses, and 0 indicating no wins. The variable summarizes the percentage of games which a player achieved their objective of winning matches.

⁵For further details on the rating system, see Appendix A.1.3.

$$percentage_{it} = \sum_{r=1}^{maxrounds_t} \frac{points_{irt}}{maxrounds_t} \in [0; 1] \quad (3)$$

The second measure is the FIDE performance rating, used to identify strong tournament results and one of the criteria for earning prestigious chess titles (FIDE, 2021). It is best understood as a counterfactual Elo rating under which a player’s win percentage against their opponents is statistically plausible. The metric accounts for opponent strength and includes an adjustment factor $dp_{it} \in [-800, 800]$ based on percentage score (Appendix A.2.7). For example, if two players achieve the same percentage score, the player who faced stronger opponents will have a higher tournament performance rating. Unlike a player’s Elo rating, which reflects a general skill level over time, the performance rating captures tournament-specific results. Higher values indicate stronger performance, and tournament performance is positively correlated with Elo ($\rho = 0.611$).

$$FIDEperformance_{it} = \frac{\sum_{r=1}^{maxrounds_t} opponentelo_{irt}}{maxrounds_t} + dp_{it} \quad , dp_{it} \in [-800; 800] \quad (4)$$

The tournament results are used as a measure of relative performance compared to all other entrants. I use the placements to indicate the cash prize winners (places 1-4, as well as best woman). Additionally, I group all the participants into quantiles to circumvent a problem that differently sized tournaments create: Placing 99th in a tournament with 150 players is different than the same place in a field of 500 players. I divide the players into quartiles, deciles, and percentiles to approximate what information players may take away after a tournament.

4.4 Descriptive Statistics

Table 1 contains summary statistics for all outcomes and key variables for the analytical sample at the between-round level. The means of the male and female subsamples are in the two left columns, and the mean and standard deviation of the full sample are additionally reported. Approximately 15% of the games are played by women, which is in line with the male-dominated nature of competitive chess.

Table 1: Summary Statistics - Game Level

	Male	Female	Full Sample			
	Mean	Mean	Mean	Median	SD	Observations
Outcome Variables:						
Average CPL	70	81	72	54	61.525	312,892
Aggressive Opening	0.268	0.257	0.267	0	0.442	203,646
Quit After Round	0.060	0.044	0.058	0	0.233	278,957
Covariates:						
Player Elo	2,538	2,318	2,506	2,500	246.746	312,892
Elo Difference	19	-71	6	77	217.064	312,892
Lost R-1	0.396	0.511	0.413	0	0.492	312,892

Note: The table presents an overview of the descriptive statistics of the relevant short run outcome variables and independent variables for the analytical sample. The analytical sample consists of games played after the third round by identified players with more than 10 moves. Means reported by gender in the two leftmost columns.

When considering the quality-of-play metric, the mean game-average centipawn loss exceeds the median by 20 centipawns, indicating a right-skewed distribution due to large errors. Skilled players usually find moves close to the computer’s suggestion, but mistakes, particularly under time pres-

sure, can produce high centipawn loss on specific moves, increasing the game-level average. Long games exhibit higher centipawn loss, consistent with more moves under mounting time pressure (see Appendix A.5). On average, women’s centipawn loss is 10 points higher, a difference that is both statistically ($p < 0.01$) and economically significant. The median centipawn loss is 54, while an average of 64 corresponds to the 60th percentile for the full sample.

Players choose an aggressive opening in 27% of games. The number of observations is smaller than the full analytical sample because only two-thirds of ECO openings were clearly classified by the expert panel (see Section 4.2.2); games with unclassified codes are omitted. Consistent with Gerdes and Gränsmark (2010), men are marginally more likely to play aggressively, doing so in 27% of games compared to 26% for women ($p < 0.01$). Without conditioning on opponent characteristics, solid openings are associated with a 51% win rate, compared to 49% for aggressive ones.

The data show that quitting is relatively rare for an online tournament: 6% of games end with a player not continuing to the next round. Quitting is defined as interrupting the tournament (see Section 4.2.3), either by leaving entirely or skipping the next round. Since tournaments end with a final round, the number of observations is slightly lower than the full analytical sample. Women quit at a roughly 25% lower rate than men ($p < 0.01$).

The Elo ratings confirm the expert nature of the sample. A rating of 2000 places a player in the top percentile of chess.com’s global pool, while the sample’s mean rating is 2506. Splitting by gender, men are on average 220 points higher rated than women ($p < 0.01$). This difference is substantial: the gap corresponds to a predicted win probability of 78% for the higher-rated player (see Equation A.1). The magnitude is consistent with over-the-board ratings, which are male-dominated at the top. Figure 3 shows the distribution of ratings by gender. Both groups’ distributions are centered around their means (Men: 2536; Women: 2316), with considerable overlap but a thicker right tail among men. Average rating differences between paired opponents are small: for men roughly 19 points, for women about 71 points lower than their opponents on average. Because Titled Tuesday uses by-points pairings, male–female matches within the same score group are therefore largely comparable in strength, even if the overall averages differ.

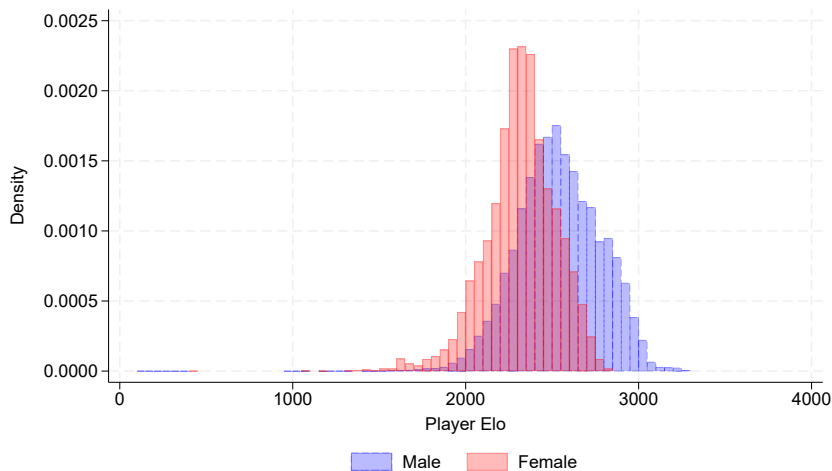


Figure 3: Elo Distribution of Players by Gender

Note: The figure presents a histogram of average within-player Elo ratings for 3,522 identified players, by gender. Men are shown in blue, women in red. Ratings are averaged across all games played by each individual over a two-year period.

The last variable is an indicator variable for whether the previous round ended in a loss. Comparing gender-specific means, 40% of male observations lost their previous round while the value is nearly 50% for female players. This is consistent what would be expected from the Elo distribution: the men’s right-tail end has higher ratings and the Elo metric predicts them to lose less than women. Given the distribution of losses, I have sufficiently large groups of losers and non-losers to compare. Across the full sample, the variable averages close to 50%⁶, retaining all observations. For matches against anonymous opponents, only the identified player is included, leaving the counterpart observation (the win corresponding to the observed loss) absent.

Table 2 presents summary statistics at the tournament level between April 20, 2020 and April 19, 2022. The panel structure allows tracking of players over time and observation of reentry behavior. The sample contains 55,875 tournament-level observations for identified players, corresponding to an average of 16 tournaments per player. Reentry choices can be recovered for 49,927 of these observations, with some data missing due to closed accounts or the final week of the sample period.⁷

Table 2: Summary Statistics - Tournament Level

	Male	Female	Full Sample		
	Mean	Mean	Mean	Median	SD
Outcome Variables:					
Weeks to Tournament Reentry	3	3	3	1	7.259
Entered Next Week	0.583	0.578	0.582	1	0.493
7 Day Sum of Blitz Games	32	17	30	11	50.030
Covariates:					
Percentage Score	0.427	0.354	0.418	0.450	0.212
FIDE Performance Measure	2,417	2,269	2,397	2,418	316.958
Tournament Elo Average	2,537	2,310	2,507	2,501	244.799
FIDE Perf. - Elo	-120	-41	-110	-54	254.270
Observations:	43,383	6,544		49,927	

Note: The table presents an overview of the descriptive statistics of the relevant long run outcome variables and independent variables for the analytical sample. The analytical sample consists of players that played in the tournament series at least twice. Means reported by gender in the two leftmost columns.

The primary tournament-level outcome is the number of days players wait to reenter the competition. Since the tournament is weekly, a 7-day wait corresponds to competing at the next possible opportunity, captured by a binary variable. The average waiting time is nearly identical for men and women (3 weeks, 23 days (men) and 24 days (women)) and not statistically different ($p = 0.376$), while the median wait is a week, reflecting that 58% of choices correspond to the next available tournament. Most reentries occur within two weeks, with a small spike at 0 days due to February 2022’s change to two tournaments per day. For a visual representation of the full distribution, including within-player averages, see Appendix Figure A.8.

The average player plays 30 blitz games per week, though men play nearly twice as many as women (Men: 32; Women: 17). A typical blitz game takes 6–8 minutes depending on moves played and time increments; using a conservative 6 minutes per game, men spend roughly 3 hours per week and women 1.5 hours on blitz play. This indicates that players invest substantial time on the platform outside the tournament, making this variable a suitable measure of engagement in a task closely related to the tournament, but less competitive.

⁶Drawn games, which account for 5% of matches, are not classified as losses.

⁷Chess.com deletes data for closed accounts, making past activity unrecoverable. Reentries after April 19, 2022 are not recorded.

As discussed in Section 4.3.3, I reconstruct several tournament-level performance metrics. The percentage score measures the share of games won, bounded between 0 and 1; a perfect score was achieved only once. Among Titled Tuesday winners, the average score is 0.890 (best woman: 0.683). On average, men finish with slightly higher scores than women (4 wins vs. 3 wins and a draw), consistent with their higher average ratings. The FIDE tournament performance metric is a counterfactual Elo rating representing the strength demonstrated in a tournament given the opponents faced. This measure confirms the sample’s high skill: both men’s and women’s tournament performance ratings remain in the top percentile of the website. On average, women underperform relative to their Elo by 44 points, while men underperform by 120 points.⁸

5 Estimating the Impact of Failure

5.1 Econometric Framework

The paper’s aim is to estimate the impact of failure on the work quality of expert chess players and study whether male and female players react differently to failure. The paper uses the multiple rounds in every tournament to study changes in three outcomes defined as the players’ quality of play, risk attitudes, and likelihood of quitting the tournament following a failure to win a previous game. To identify the effect of failure (losing a match), I use a simple TWFE (2x2 differences-in-differences) framework. The treatment effect δ is the difference in expected outcomes of players that lost the previous round (treated) and those that did not (control). Equation 5 therefore describes the average treatment effect on the treated (ATT).

$$\delta = \left[E(Y_{treat}|After) - E(Y_{treat}|Before) \right] - \left[E(Y_{control}|After) - E(Y_{control}|Before) \right] \quad (5)$$

In order for this estimation to be approximated with the sample means, a (conditional) parallel trends assumption must hold. Rewriting this equation in potential outcomes notation makes the necessary underlying assumption clear: the treatment group (those that lost in a given round) would have had the same performance and trajectory as the control group (those that did not lose in a given round). Expanding and rewriting the initial equation shows that the estimates may be biased if this assumption fails. Subscripts denote the the player group and superscripts indicate the potential outcome.

$$\delta = \left[E(Y_{treat}^{lost}|After) - E(Y_{treat}^{won}|After) \right] + \left[E(Y_{treat}^{won}|After) - E(Y_{treat}^{won}|Before) - E(Y_{control}^{won}|After) - E(Y_{control}^{won}|Before) \right] \quad (6)$$

The first term $E(Y_{treat}^{lost}|After) - E(Y_{treat}^{won}|After)$ is a counterfactual term, as it involves the loser’s groups outcomes had they won, which cannot be observed. The subsequent estimations will use a multivariate specification to construct this counterfactual from the control group. The second term describes the bias in the estimation if the parallel trends assumption does not hold. The implications of the estimation are that any differences in outcomes between the losers and winners are driven by the loss itself and not anything else, conditional on a vector of observable variables X .

⁸Conditional on completing the tournament, both genders overperform: women by 42 points, men by 13 points.

$$E(Y_{treat}^{lost}|X) = E(Y_{control}^{lost}|X) \quad (7)$$

$$E(Y_{treat}^{won}|X) = E(Y_{control}^{won}|X) \quad (8)$$

Provided that the losers and winners of a round are similar in characteristics and the losers would have otherwise had a similar performance to the winners, the regression specification in Equation 11 yields the causal estimate for the impact of losing on subsequent tournament behavior.

This paper’s goal is to study gender differences in reaction to failure. The empirical strategy outlined above does not require women to behave identically to men or be comparable to men. The empirical strategy compares the within-group outcomes of those who lost to those who won. The gender-specific ATT may be formalized as:

$$\begin{aligned} \delta_m = & \left[E(Y_{treat}|After, Female = 0) - E(Y_{treat}|Before, Female = 0) \right] \\ & - \left[E(Y_{control}|After, Female = 0) - E(Y_{control}|Before, Female = 0) \right] \quad (9) \end{aligned}$$

$$\begin{aligned} \delta_f = & \left[E(Y_{treat}|After, Female = 1) - E(Y_{treat}|Before, Female = 1) \right] \\ & - \left[E(Y_{control}|After, Female = 1) - E(Y_{control}|Before, Female = 1) \right] \quad (10) \end{aligned}$$

In particular, there is no requirement that $\delta_m = \delta_f$. As before, the underlying assumption is that any differences in outcomes are being driven by losses. Women and men who lost would otherwise behave like the women and men who won.

5.2 Regression Specification

5.2.1 Within Tournament Reaction to Failure

The baseline specification takes on the form:

$$Y_{irt} = \alpha_i + \sum_{k=1}^3 \delta_k \text{Loss}_{i,r-k} + \sum_{k=1}^3 \gamma_k \text{Loss}_{i,r-k} \times \text{Female}_i + \Theta X_{irt} + \text{round}_r + \text{tournament}_t + \epsilon_{irt} \quad (11)$$

For the within-tournament regressions, the outcome Y_{irt} is defined on a game-level for a player i in round r of tournament t . I focus on three game-level outcomes to study behavioral response to failure, as outlined in Section 4.2: cognitive performance, whether the chosen opening is classified as “aggressive”, and if players quit the tournament.

The variable $\text{Loss}_{i,r-k,t}$ is an indicator variable which equals one if a previous round $r - k$ ended in a loss. To capture possible differential reactions between men and women, the variable $\text{Loss}_{i,r-k,t}$ is interacted with the indicator Female_i , which takes on a value of one if a player is female. The coefficients of interest are δ_k , which quantifies the impact of losing for men, and γ_k which captures the differential behavioral response of women compared to men. I include up to three previous results in the main specification.

The vector of variables X_{irt} encompasses a variety of match controls. To control for the differing skill levels in the sample, the player’s Elo rating is included. A significant source of variation in chess stems from the opponent. Chess.com randomly pairs players within a points bracket. I include all information that a player would have about their opponent: their rating difference, the opponent’s title and profile picture. I also control for piece color to capture any behavioral differences between White and Black. Additional player-specific controls for age and its square are included. Every specification includes person fixed effects, round fixed effects, and tournament fixed effects.

For the coefficients δ_k on the $Lost_{i,r-k,t}$ indicators (and γ_k on the $Lost_{i,r-k,t} \times Female_i$ interactions) to have a causal interpretation, the control and treatment groups must be comparable. The main concern is that weaker players are more likely to lose, making the “treatment” non-random. However, “weaker” is defined only relative to a randomly allocated opponent within a points pool, and even higher-rated players lose frequently: the stronger player wins 63% of the time. Furthermore, most participants end tournaments with a percentage score between 0.4 and 0.6 (Figure A.2). My identification exploits variation in *when* players lose rather than *if* they lose. As discussed in Section 4.3.1, Chess.com Elo ratings provide a precise numerical measure of player skill, ensuring that loss timing is largely independent of baseline ability. Because most players lose at least one game during a tournament, there is substantial overlap in outcomes across the sample. A second concern is that players might deliberately lose to influence future pairings. This is unlikely given the tournament structure: rankings are determined by total points, all rounds are played, and there is no consolation bracket. All games are rated and deliberate losses reduce Elo ratings, which are costly for players given their importance in identity and recognition. To further account for differences in observable characteristics, I include controls for factors that may independently affect performance, risk-taking, and quitting.

I include player fixed effects to absorb time-invariant individual characteristics such as baseline playing strength, experience, or preferred style of play. This ensures that my estimates are identified from within-player variation across rounds rather than between-player differences. Round fixed effects capture systematic variation as the tournament progresses: since a typical event lasts about three hours, players may become fatigued over time, which could otherwise bias estimates of post-loss performance. Tournament fixed effects control for common shocks in a given event, such as chess.com server instability, or changes in the competitive pool if certain top players participate (Bilen and Matros, 2023). Together, these sets of fixed effects remove the most likely sources of confounding variation at the player, round, and tournament level. Finally, I control for age, which may influence both cognitive performance (Strittmatter et al., 2020).

The controls include all factors directly relevant for chess play that are observable to participants such as information about their opponent and a precise measure of chess skill. In other words, I control for everything that the players see during a game. A potential external factor remaining is internet connectivity. Poor connection could hamper performance through move latency or misclicks and influence quitting behavior. However, most players in the sample are highly experienced, with Elo ratings around 2500, well above the 1800 starting point for new accounts. They are likely aware of their connection quality and its potential impact. Players with consistently poor connectivity are unlikely to enter a competitive speed chess tournament. Additionally, since the analytical sample starts at round 4, early quits due to connectivity are already excluded. Provided that these variables sufficiently account for potential differences between the treatment and control group, the parameters δ_k and γ_k represent the causal effect of losing.

5.2.2 Between Tournament Reaction to Outcome

To estimate whether there are gender differences in willingness to compete in the long run, I estimate how many days are needed until individuals are willing to reenter into the competition. The baseline specification is a linear model of the form:

$$Days_{it} = \alpha_i + \delta TourPerf_{i,t-1} + \gamma TourPerf_{i,t-1} \times Female_i + \theta X_{it} + tournament_t + \epsilon_{it} \quad (12)$$

The outcome is the days elapsed until a person reenters the tournament. The variable $TourPerf_{i,t-1}$ summarizes a player’s most recent tournament performance. I use three different types of performance metrics, as detailed in Section 4.3.3: the percentage score, the tournament performance rating, and the placement quartile, with the bottom quartile as the reference category. The coefficient δ describes how many days the average man needs until he is willing to reenter into competition for a given performance. The differential effect for women is captured by the coefficient γ . Each regression incorporates player fixed effects to control for the players’ baseline preferences for competition and tournament fixed effects. Additional controls are included in the vector X_{it} for the players’ rating at the start of the tournament, their age, and whether they previously won a cash prize, which may influence their incentive to reenter.

Additionally, I examine how account activity changes following tournament participation. Because the tournament is played in a speed chess format that emphasizes a trade-off between calculation and time management, playing more non-tournament blitz games on chess.com indicates additional practice.

$$Games_{it} = \alpha_i + \delta TourPerf_{i,t-1} + \gamma TourPerf_{i,t-1} \times Female_i + \theta X_{it} + tournament_t + \epsilon_{it} \quad (13)$$

The outcome variable $Games_{it}$ is the number of non-tournament rated blitz games played on chess.com in the 7 days following a tournament participation. If a player plays back-to-back weeks, the variable captures the total number of speed chess games played between these two tournaments. If a player skips the next possible tournament, the variable reflects the number of games they played until the next competition opportunity. The rest of the specification follows Equation 12. The coefficient δ and γ describe the relationship between past tournament performance and subsequent engagement with (online) speed chess for men (δ) and possible differential effects for women (γ). The sign of these two effects is theoretically ambiguous: a bad performance may induce players to increase their speed chess activity to practice the format ($\delta, \gamma > 0$) or players may face a higher emotional cost engaging in an activity where they perform worse than their peers ($\delta, \gamma < 0$).

The model in Equation 13 should be regarded as descriptive. A key concern is reverse causality: players may perform worse *because* they spend less time on speed chess. If differences were driven by baseline skill - top players naturally investing more time than lower-ranked players, then player fixed effects would account for this. Controlling for Elo ratings provides an additional check. Together, these controls should largely account for differences in activity across playing strength and within-player. Playing additional speed chess games is one of many ways players invest in their chess skills, this variable only captures online activity; offline practice cannot be measured. Consequently, any patterns identified are descriptive and only represent a subset of a player’s total chess activity.

6 Results

The results center around three outcomes at a within-tournament level following model specified in Equation 11.

6.1 Losses and Cognitive Performance

The players' goal is to play with as few mistakes as possible. The dependent variable is a player's average centipawn loss, which captures the difference in advantage between the computer's optimal move and the move the player actually made. If a player makes no mistakes that their opponent can capitalize on, this implies a low average centipawn loss. When examining how previous rounds affect average centipawn loss, a positive coefficient implies a decreased quality of play while a negative coefficient indicates more precise play and fewer mistakes.

Table 3: Effect of Previous Rounds on Average Centipawn Loss

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.242 (0.258)	0.476* (0.277)	1.086*** (0.363)	1.251*** (0.381)	-0.552** (0.266)	-0.436 (0.292)
Lost R-2	0.948*** (0.249)	0.970*** (0.269)	1.937*** (0.403)	2.132*** (0.428)	-0.018 (0.274)	-0.218 (0.300)
Lost R-3	0.451* (0.248)	0.500* (0.265)	0.708* (0.377)	0.929** (0.393)	-0.051 (0.267)	-0.188 (0.295)
Lost R-1 x Female		-1.533** (0.655)		-1.399 (1.182)		-0.665 (0.664)
Lost R-2 x Female		-0.118 (0.664)		-1.711 (1.252)		1.126 (0.687)
Lost R-3 x Female		-0.305 (0.711)		-2.118 (1.361)		0.715 (0.675)
Observations	312,818	312,818	312,818	312,818	312,818	312,818
R ²	0.048	0.048	0.048	0.048	0.048	0.048

Note: Dependent variable is the average centipawn loss per move. A positive value implies deviation from the computer-preferred move resulting in an avoidable disadvantage. All regressions control for the player's own rating, the rating difference to their opponent, their age, their opponent's title and the gender signaled in the opponent's profile picture. All regressions use player, tournament and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Past results affect the current round's playing precision. For losses in previous rounds, the coefficients are generally positive, suggesting a measurable drop in play precision of roughly 0.2–0.9 centipawns per move. Losing two rounds ago increases the average deviation from the computer's preferred move by about one centipawn, a significant deterioration in performance. Column (2) of Table 3, which includes interactions with a female indicator, highlights gender differences in responses to losses. On average, men's performance declines after a negative outcome by approximately 0.8% of the mean (approx. half a centipawn), whereas women's performance improves relative to men in the subsequent round. These differential responses, captured by the interaction terms, create a gap in playing quality of roughly one centipawn between men and women, equivalent to 1.5% of the mean.

Upset losses, losing against a lower-rated opponent, are particularly salient, representing a game that should have been won. Section 4.3.1 highlighted that the higher rated player is the favorite

at match begin and wins their games 63% of the time; my results indicate that an upset loss has at least twice the detrimental effect of any loss (see Table 3, Columns(3) and (4)). Both men’s and women’s cognitive performance declines after upset losses. By contrast, expected losses, from matches against higher-rated opponents, do not meaningfully affect performance, highlighting that the disruptive effect is specific to unexpected defeats (see Table 3, Columns(5) and (6)). I do not find robust evidence of gender differences in processing upset losses. While the point estimates for the female interaction terms are negative, they are imprecisely estimated and insignificant, reflecting the fact that upset losses by women account for only about 1% of all games in the sample. When honing in on lost matches where women were the higher-rated favorite, the loss appears to have a similar “sting” for both men and women.

This change in cognitive performance can have implications for subsequent match outcomes. Following a loss, men’s playing precision decreases by approximately half a centipawn per move. Women’s playing precision improves by one centipawn. This change in playing quality translates to women improving by one percentile from the median centipawn loss.⁹ An average centipawn loss increase of 1 (mistakes computationally equal to $\frac{1}{100}$ of a pawn per move) is associated with a 0.1pp reduction in win probability. A back-of-the-envelope calculation then suggests that men’s decline in cognitive performance by a loss reduces their predicted win probability by 0.05pp. Women’s stronger cognitive performance would translate to a increase of the predicted win probability by 0.1pp.

6.2 Losses and Opening Choice

When studying opening aggressiveness, Table 4 shows that male players respond to previous losses by taking on more risk in the opening phase of the game while women did not change their baseline propensity for aggressive play. I use the same classification dataset as [Gerdes and Gränsmark \(2010\)](#) to study how the willingness to take risk through aggressive play changes with tournament progression. I only include games that resulted in a clear classification of “solid” or “aggressive”. Positive coefficients indicate a higher likelihood of risk-taking.

Losing the previous round increases the probability of choosing an aggressive opening and taking on more risk early in the game by 1 percentage point (Table 4, Column 1). The additional backwards-looking terms suggest that outcomes from earlier rounds influence risk-taking in the current match. Losing two rounds prior still increases the likelihood of selecting an aggressive opening. As expected, more recent losses have a larger effect size than those further in the past.

To examine gender differences in reactions to loss, I interact the three lagged round results with a female dummy. The lagged results without an interaction should be interpreted as the average effect for men, whereas the average effect for women is the sum of the lagged and interacted coefficients. Column (2) in Table 4 provides weak evidence that the gender gap in risk-taking is partly driven differential responses to losing: men increase their likelihood of choosing an aggressive opening following a loss, whereas women’s behavior remains largely unchanged. Because this analysis exploits within-player variation, these results do not simply reflect the fact that men generally select aggressive openings more often; instead, they summarize how each player adjusts their strategy in response to prior outcomes. Men are more likely to take additional risk by choosing an aggressive opening in subsequent rounds. In short, losing a round leads men, but not women, to increase risk-taking in the opening.

⁹The results are robust to using censored centipawn loss data.

Table 4: Effect of Previous Rounds on Opening Riskiness

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.009*** (0.002)	0.011*** (0.002)	0.016*** (0.003)	0.019*** (0.004)	0.001 (0.002)	0.001 (0.002)
Lost R-2	0.008*** (0.002)	0.009*** (0.002)	0.004 (0.003)	0.004 (0.003)	0.005** (0.002)	0.005** (0.002)
Lost R-3	0.002 (0.002)	0.002 (0.002)	0.000 (0.003)	0.003 (0.003)	-0.001 (0.002)	-0.002 (0.002)
Lost R-1 x Female		-0.012** (0.006)		-0.019** (0.010)		-0.003 (0.006)
Lost R-2 x Female		-0.006 (0.006)		0.001 (0.011)		-0.002 (0.006)
Lost R-3 x Female		-0.002 (0.005)		-0.027*** (0.010)		0.008 (0.006)
Observations	203,549	203,549	203,549	203,549	203,549	203,549
R ²	0.146	0.146	0.146	0.146	0.146	0.146

Note: Dependent variable is a binary variable indicating whether the opening employed is deemed “aggressive”. All regressions include controls for the player’s own rating, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions include player, tournament, and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The dataset’s detail allows me to directly examine how behavior changes with different types of losses and the influence of expectations. Unexpected losses are defined as losing to an opponent with a lower Elo rating. Column (3) of Table 4 shows that these losses have a larger impact: players who suffered an unexpected loss in the previous round are close to two percentage points more likely to choose an aggressive (riskier) opening compared to those who did not. This effect is highly statistically significant and nearly 145% the magnitude of a general loss (compare Table 4 column (1)). Interacting upset loss terms with a female dummy mirrors the patterns observed for any loss: men respond more aggressively following unexpected losses, while women’s behavior remains unchanged, as the sum of the coefficients is close to zero. Losses against stronger opponents (expected losses) show no such effect. Taken together, these results demonstrate that men’s increased risk-taking is specifically triggered by unexpected defeats.

6.3 Losses and Tournament Abandonment

The Titled Tuesday format involves multiple rounds played in immediate succession. Players can quit mid-tournament by closing the browser, for example to avoid further disappointment. To capture this behavior, I verify whether each player participates in round $r + 1$ (when r is not the final round). This provides a measure of both if and when a player dropped out.

Table 5 summarizes the effect of losses on quitting behavior. Columns (1) and (2) highlight that quitting is primarily an immediate reaction to the last game. Losing a round is associated with a 5 percentage points higher likelihood of quitting the tournament afterwards. The prior round results continue influence quitting choices, but to a far smaller degree, approximately 1pp. Immediately after losing, men are nearly 13pp more likely to quit the tournament compared to those that won or drew. While losing also increases the quit likelihood for women, it does so by a far lesser margin of around 8pp, two-thirds of the men’s effect size. The results suggest that men and women differ in their immediate reaction to a loss, but place similar weight on earlier rounds.

Table 5: Effect of Previous Outcomes on Quitting Behavior

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.122*** (0.003)	0.129*** (0.004)	0.171*** (0.005)	0.176*** (0.006)	0.049*** (0.002)	0.051*** (0.002)
Lost R-2	0.020*** (0.001)	0.021*** (0.001)	0.024*** (0.002)	0.023*** (0.002)	0.007*** (0.001)	0.007*** (0.001)
Lost R-3	0.012*** (0.001)	0.011*** (0.001)	0.008*** (0.002)	0.008*** (0.002)	0.004*** (0.001)	0.003** (0.001)
Lost R-1 x Female		-0.049*** (0.006)		-0.050*** (0.011)		-0.012*** (0.004)
Lost R-2 x Female		-0.004 (0.003)		0.006 (0.005)		0.002 (0.003)
Lost R-3 x Female		0.002 (0.003)		0.001 (0.005)		0.006** (0.003)
Observations	278,887	278,887	278,887	278,887	278,887	278,887
R ²	0.128	0.129	0.132	0.132	0.077	0.077

Note: Dependent variable is a binary variable indicating whether a player interrupted their tournament play. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player, tournament, and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Columns (3) and (4) of Table 5 isolate the effect of upset losses, when players lose to a lower-rated opponent. These losses are particularly salient, producing a stronger quitting response than average losses. Men remain more likely to quit following an upset loss, while women’s response is smaller, about two-thirds of the male effect. The magnitude of an upset loss is approximately 1.4 times larger than a general loss for both genders. By contrast, expected losses have a smaller impact on quitting behavior (Columns (5) and (6)). Losing against a higher-rated opponent increases the quitting probability 4–5pp, about 40% of the effect of an average loss. Overall, quitting is primarily driven by defeats that a player should have won. Both men and women are more likely to quit following a loss, but the effect on men is consistently larger, up to 1.5 times the effect on women. Quitting behavior mirrors other behavioral responses: losses that “should not” happen trigger stronger immediate reactions, and men are consistently more sensitive than women.

6.4 Placement and Time between Tournaments

The Titled Tuesday tournament is a weekly recurring series. Players can therefore decide how often they wish to compete in the series, providing an ideal setting to study longer-run tournament reentry behavior. There is a possible tournament every week, so a waiting time of a single week corresponds to reentering into competition at the first available opportunity. Table 6 summarizes the results using a linear regression with the weeks until reentry as a dependent variable. This pattern holds across a variety of performance measures. Strong performers reenter the tournament series faster, while weaker performers wait longer.

The percentage score measures absolute success in the tournament, capturing the share of matches won and bounded between 0 and 1. As shown in Column (1) of Table 6, players who won every game are expected to reenter approximately three weeks sooner than those who did not win any

Table 6: Tournament Performance and Weeks Waited Until Reentry

	(1)	(2)	(3)	(4)	(5)	(6)
Percentage Score	-3.399***	-3.138***				
	(0.229)	(0.238)				
Percentage Score \times Female		-2.617***				
		(0.799)				
FIDE Perf. – Elo (in 100s)			-0.210***	-0.201***		
			(0.016)	(0.017)		
FIDE Perf. – Elo (in 100s) \times Female				-0.085		
				(0.052)		
Placement Quartile						
1 st Quartile					-1.705***	-1.673***
					(0.129)	(0.137)
2 nd Quartile					-1.447***	-1.433***
					(0.116)	(0.126)
3 rd Quartile					-1.040***	-1.006***
					(0.111)	(0.122)
Placement Quartile \times Female						
1 st Quartile \times Female						-0.311
						(0.337)
2 nd Quartile \times Female						-0.092
						(0.295)
3 rd Quartile \times Female						-0.210
						(0.285)
Observations	52,126	52,126	52,126	52,126	52,126	52,126
Adjusted R ²	0.222	0.222	0.221	0.221	0.221	0.221

Note: Dependent variable is how many weeks elapsed to the next tournament entry. All regressions include additional controls for the player’s rating at the day of the reentry choice and indicators whether they won a cash prize in their last tournament appearance. All regressions include player and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

games. Notably, this effect is slightly stronger for women, where particularly strong performances are associated with an additional 2 week waiting time reduction.

A player’s Elo rating predicts the expected level of performance against their opponents, so outcomes above or below this expectation indicate over- or underperformance relative to what their rating would predict. Columns (3) and (4) of Table 6 report the effects of this relative performance, measured as the difference between a player’s FIDE performance and their Elo rating. A positive value indicates a tournament performance that exceeded the performance predicted by their rating (and vice versa for a negative value). Players who overperformed relative to their rating in their previous tournament tend to reenter sooner. For example, a 100-point overperformance corresponds to entering the next tournament roughly 1.4 days earlier. There is no evidence that tournament reentry patterns, conditional on over- or underperformance relative to rating, differ between men and women.

While percentage scores and rating-based performance measures provide useful information, tournament success is most often judged by placement relative to other competitors. Columns (5) and (6) report expected waiting time until tournament reentry, conditional on a player’s placement in the last tournament. To capture the impact of tournament placement independent of the tournament size, I use quartiles with the bottom quartile as a base. As with previous measures of absolute and relative performance, higher placement predicts faster reentry. Players who finished in the top 25% reenter about 1.7 weeks (11 days) sooner compared to the bottom 25%. Players in the second quartile reenter about 1.4 weeks (10 days) earlier than those in the bottom quartile. Including gender interactions in column (6) does not alter the estimates, and all female interactions are insignificant. These results indicate that men and women respond similarly to tournament placement.

Across both subjective and relative performance measures, strong performers consistently reenter

the tournament faster than weaker performers. Overall, the results present compelling evidence that tournament reentry is systematically influenced by past performance. In particular, the result is consistent with rational belief-updating: strong performers may feel emboldened by their performance and have reason to believe that they will perform similarly next time. These results are robust to using a Poisson specification (see Table B.12) as well as a binary specification on whether players enter a week after their last performance, i.e. the next possible tournament (see Table B.13). Notably, there are no robust gender differences between men and women in how they interpret the results and their reentry behavior converges.

6.5 Placement and Additional Practice

Beyond studying how previous performance influences tournament reentry, this section examines how success and failure affect practice behavior. Non-tournament blitz games in the week following a tournament are used as a measure of practice activity and are regressed on previous tournament performance (Equation 13). The dependent variable is the number of blitz games played in a week: a positive coefficient indicates that a factor is associated with increased practice, while a negative coefficient indicates the opposite.

Table 7: Tournament Performance and Subsequent Platform Use

	(1)	(2)	(3)	(4)	(5)	(6)
Percentage Score	10.197***	10.114***				
	(1.284)	(1.353)				
Percentage Score \times Female		0.860				
		(3.795)				
FIDE Perf. – Elo (in 100s)			0.554***	0.556***		
			(0.081)	(0.081)		
FIDE Perf. – Elo (in 100s) \times Female				-0.008		
				(0.009)		
Placement Quartile						
1 st Quartile					4.417***	4.286***
					(0.825)	(0.875)
2 nd Quartile					3.689***	3.507***
					(0.684)	(0.762)
3 rd Quartile					2.674***	2.546***
					(0.623)	(0.713)
Placement Quartile \times Female						
1 st Quartile \times Female						1.075
						(1.785)
2 nd Quartile \times Female						1.315
						(1.402)
3 rd Quartile \times Female						0.863
						(1.041)
Observations	51,974	51,974	51,974	51,974	51,974	51,974
Adjusted R ²	0.459	0.459	0.458	0.458	0.458	0.458

Note: Dependent variable is the number of rated blitz played on chess.com in the week after a tournament participation. All regressions include controls for the player’s last rating and use player and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column (1) in Table 7 uses the percentage score as a performance metric. The percentage score is the player’s total score after all rounds divided by the number of rounds and is in the interval between 0 and 1. The coefficient suggests that a perfect score (only won games) is associated with a total of 10 additional games in the following week compared to only-losers. Incorporating a gender interaction in column (2) does not change the interpretation: the male point estimate remains unchanged and the female interaction coefficient is insignificant. In terms of time investment, a blitz chess game takes approximately 6 minutes. Interpreted as practice behavior, this roughly corresponds to an extra hour of practice at tournament blitz time controls, regardless of gender.

I can also measure tournament performance relative to a player’s expected level, as predicted by their Elo rating (see Section 4.3.3). Column (3) in Table 7 shows that a 100-point increase in relative performance is associated with roughly half an additional blitz game in the following week. Conversely, underperformance is associated with fewer games or practice in the following week. The inclusion of a gender interaction in column (4) does not change the interpretation: the main effect for men remains, and the female interaction is not statistically significant. Interpreted as practice behavior, these results suggest that both men and women increase their post-tournament practice similarly following strong performances.

Column (3) of Table 7 examines placement quartiles, using the bottom 25% as the reference group. Players who finished higher in the standings consistently played more blitz games in the following week, with the number of games increasing across quartiles. On average, those in the top 25% played about 4 additional games in the following week relative to the bottom quartile. This pattern mirrors the reentry results: strong performers invest more time in chess practice outside the tournament, just as they return to competition sooner. As before, gender interactions are insignificant, indicating that men and women interpret their tournament results similarly and adjust their practice behavior in comparable ways.

Across all performance measures, stronger tournament results are consistently associated with greater subsequent practice, while weaker results correspond to significantly less investment of time. Because the analysis exploits within-player variation across multiple tournaments, these effects reflect how the same individual adjusts practice behavior after better or worse outcomes, rather than differences across players. The patterns closely parallel the reentry results: players appear to use tournament performance as a signal that guides both their competitive participation and their training effort. However, these findings are correlational and capture only one dimension of chess activity—blitz games played online—so they should be interpreted as evidence of a robust association rather than causal proof. Finally, as with reentry behavior, there is no evidence of systematic gender differences: men and women respond to tournament outcomes in similar ways when it comes to practice.

7 Discussion

I have presented evidence of differing behavioral responses by men and women following failure. As discussed in Section 5, the parameters of interest are causally identified if the treated (lost a match) and control group (won a match) would have progressed similarly had it not been for the failure. In this setting, this means that players who land in the same opponent pool via different arms (coming off a win versus a loss, compare Figure 1) are otherwise similar.

The short-term results show that men’s performance is more volatile than women’s. Men’s short-term reactions to failure, including decreased precision, increased risk-taking, and higher quit rates, suggest an emotionally-driven response to setbacks, particularly unexpected losses. Women, in contrast, maintain or even improve their performance, consistent with stronger resilience and stability under negative feedback. Similar patterns have been observed in other high-stakes competitive settings. In the heavily male-dominated mutual fund industry (Aggarwal and Boyson, 2016), female managers show more stable performance and lower risk profiles over time, yet maintain comparable returns to men (Niessen-Ruenzi and Ruenzi, 2019). In educational settings, narrowly failing the first university exam reduces men’s likelihood of finishing their degree, whereas women’s grades on

the second attempt exceed men’s (Mueller, 2023; Kang et al., 2024). These parallels indicate that the behavioral differences observed in chess reflect a broader pattern in competitive environments.

Turning to longer-run reentry behavior, I find no gender differences in the willingness to reselect into competition conditional on past performance feedback. Strong performers - whether male or female - are faster to reenter the tournament series, echoing results found in Wozniak (2012). These results initially appear to contrast the “common” finding that women refrain from competing at greater rates than men (Alnamlah and Gravert, 2020; Markowsky and Beblo, 2022; Niederle and Vesterlund, 2007); however, my findings align with experimental evidence that regular feedback helps close the gender gap in willingness to compete (Buser, 2016; Ertac and Szetes, 2011). The institutional features of the Titled Tuesday tournament make feedback unavoidable. Players possess near-perfect information: they have information about their ability relative to others, on the game results (and how the result came to be), but also about their opponent. During a game, rating information is visible to both players. Winning against a higher-rated player or losing against a lower-rated player offers objective performance feedback to the participants. When the tournament is finished, players should be able to draw accurate conclusions about their own ability relative to the competitor pool. The finding that stronger players tend to be the fastest to reenter is consistent with rational belief updating based on past performance. Because they are being evaluated along the same lines, men and women appear to place the same weight on past results. The week-long interval between tournaments allows players to step back from short-term emotional reactions. While we cannot directly observe their thought processes, the pattern of longer-term reentry behavior is consistent with players using their previous tournament performance as a reference point: they appear to assess the likelihood of success in the upcoming tournament based on prior results, weighing whether they are likely to perform well or risk another disappointing outcome. In this sense, longer-term reentry choices reflect considered evaluations of skill and expected outcomes, rather than immediate, emotionally-driven reactions.

A natural concern is how far the findings extend beyond chess. The results should not be interpreted as representative of the general population. Instead, they are most relevant to women in male-dominated, competitive fields such as finance, STEM, or academia, where advancement depends on sustained performance in cognitively demanding and non-routine tasks. The women in my sample have already broken into the competitive circle by earning their titles, comparable to women who have entered professional pipelines and are in the process of climbing to the top. Chess is a useful lens in this context because it allows unusually detailed measurement of short-term risk-taking behavior and mistake magnitude in a cognitively demanding task (Bertoni et al., 2015), similar to Autor and Handel’s (2013) characterization of managerial work. Economists have previously relied on chess to shed light on workplace outcomes: Künn et al. (2022, 2023) use chess to study the determinants of work quality, and De Sousa and Niederle (2022) analyze affirmative action quotas in chess to speculate about their transferability to academic settings. Psychology research suggests that intelligence does not strongly predict chess skill level (Bilalić et al., 2007a; Waters et al., 2002; Grabner et al., 2007; Unterrainer et al., 2006). Certain personality traits, particularly those related to impulse control, are positively associated with chess ability and more prevalent among players than in the general population (Bilalić et al., 2007b; Blanch and Llaveria, 2021; Grabner et al., 2007). Male chess players’ personalities largely mirror population norms, while female players tend to be slightly more achievement-oriented (Vollstädt-Klein et al., 2010). Economists have studied chess players under the hypothesis that their cognitive training for move calculation might translate into superior performance in other economic games. In a beauty contest, Bühren and Frank (2012) find chess players fail to outperform previous student samples.

Similar findings have been echoed by [Levitt et al. \(2011\)](#) in the centipede game.¹⁰ Overall, chess players' performance in economic games outside of their domain does not systematically differ from standard subject pools. Together, this evidence supports interpreting my findings as most relevant to competitive environments where women have already entered the field, rather than as a general reflection of the broader population.

My findings suggest that conditional on performance feedback, men and women that were initially willing to compete do not react differently to losses in the long run. In the context of job applications and promotions, this suggests that career-oriented men and women recover from setbacks at similar rates in the longer run. The reaction to failure is not a channel through which the gender leadership gap widens. However, chess is a reasonably controlled setting: game and tournament results are objective, transparent, and nearly impossible to misinterpret. In contrast, feedback in professional or academic settings is noisier. Men and women receive different types of feedback ([Dupas et al., 2021](#); [Chakraborty and Serra, 2024](#)), recommendation letters ([Eberhardt et al., 2023](#)), and credit for the same performance ([Benson et al., 2022](#); [Sarsons et al., 2021](#)). Even if men and women respond the same way to feedback, differences in feedback quality, such as women being told they ranked in the third quartile when a comparable man is told he ranked in the second, could lead women to make different choices, creating gender gaps in competitive settings despite similar underlying willingness to compete.

Differences in dealing with negative outcomes do not appear to affect long-term willingness to reengage in competition. Yet, female representation still declines at the top-most positions: even among women willing to compete, fewer reach the highest levels ([Gneezy et al., 2017](#); [Bertrand and Hallock, 2001](#); [Goldin, 2014](#)). If women do not systematically choose to opt out of the competition, this suggests that other mechanisms preventing a proportional representation are at play, such as glass ceiling effects or discrimination. My field setting highlights the importance of information signals in sustaining competitive engagement and supports findings from laboratory experiments ([Ertac and Szetes, 2011](#)). Men and women in chess receive the same performance feedback, and reentry behavior aligns with expectations formed from prior outcomes: strong performers reenter sooner. When feedback is clear, observable, and minimally ambiguous, there is less room for (unconscious) gendered bias, pointing to the potential importance of equitable signal quality in competitive environments.

8 Conclusion

This paper uses 286,140 chess games of 3,523 title-holding players in a popular tournament series as a naturally occurring setting to examine how failure affects short-run performance and longer-run willingness to compete, and whether there are differences along gender lines. Following losses, men take on more early-game risks than women. Comparing opening choices shows that men are more likely to choose aggressive openings than women following a loss, conditional on a very rich set of player and opponent characteristics. Similarly, men are more likely to quit a tournament after a loss than women, suggesting that women are better able to recover in the immediate aftermath of a loss. These results are echoed in the analysis of playing accuracy. Comparing the moves made by a chess player against those preferred by a powerful chess engine reveals that women's cognitive performance improves after a loss, while men's declines. Overall, the results suggest that in the context of experienced professionals, women are more resilient to negative outcomes compared to men and improve their performance following adverse outcomes.

¹⁰A prior paper by [Palacios-Huerta and Volij \(2009\)](#) finds opposite results for the centipede game, but could not be replicated in a similar field setting.

An analysis of account activity after tournaments shows that strong tournament performances are linked to increased play. Both men and women significantly increase their non-tournament play on the platform after a successful performance, with no evidence for gender differences at the intensive margin. For men, finishing in the top quartile relative to the bottom is associated with an increase in weekly games equivalent to 13% of the mean (4 games). For women, a top quartile finish is associated with a 25% increase of weekly games relative to the mean (4 games). Overall, the results point to success having a similarly encouraging effect to investing time into improving skills for men and women.

Exploiting the dataset's two-year timespan, I compare the number of weeks elapsed between tournament appearances for men and women. The data shows that on average three weeks go by between tournaments and that there are no significant gender gaps in tournament re-entry, conditional on past tournament performance. Strong entrants are quicker to reenter into an event compared to those that achieved a low number of points. A possible mechanism explaining the missing gap in the field could be the immediate and objective feedback mechanisms that chess provides, giving transparent individual performance feedback as well as relative performance feedback. These findings suggest that in environments with transparent and immediate feedback, gender differences in long-term engagement may be minimal.

References

- Acher, M. and Esnault, F. (2016). Large-scale Analysis of Chess Games with Chess Engines: A Preliminary Report. Technical Report 479, University of Rennes, Rennes, France.
- Aggarwal, R. and Boyson, N. M. (2016). The Performance of Female Hedge Fund Managers. *Review of Financial Economics*, 29:23–36.
- Ahmed, I., Sultana, I., Paul, S. K., and Azeem, A. (2013). Employee Performance Evaluation: A Fuzzy Approach. *International Journal of Productivity and Performance Management*, 62(7):718–734.
- Alnamlah, M. and Gravert, C. (2020). She Could Not Agree More: The Role of Failure Attribution in Shaping the Gender Gap in Competition Persistence. CEBI working paper series 20-25, University of Copenhagen. Department of Economics. The Center for Economic Behavior and Inequality (CEBI).
- Autor, D. H. and Handel, M. J. (2013). Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, 31(2):S59–S96.
- Backus, P., Cubel, M., Guid, M., Sanchez-Pages, S., and Manas, E. (2023). Gender, Competition, and Performance: Evidence from Chess Players. *Quantitative Economics*, 14(1):349–380.
- Bapna, S., Benson, A., and Funk, R. (2021). Rejection Communication and Women’s Job-Search Persistence. Working Paper.
- Benson, A., Li, D., and Shue, K. (2022). “Potential” and the Gender Promotion Gap. Technical report.
- Berlin, N. and Dargnies, M.-P. (2016). Gender Differences in Reactions to Feedback and Willingness to Compete. *Journal of Economic Behavior & Organization*, 130:320–336.
- Bertoni, M., Brunello, G., and Rocco, L. (2015). Selection and the Age – Productivity Profile. Evidence from Chess Players. *Journal of Economic Behavior & Organization*, 110(February):45–58.
- Bertrand, M. and Hallock, K. F. (2001). The Gender Gap in Top Corporate Jobs. *Industrial and Labor Relations Review*, 55(1):3–21.
- Bilalić, M., McLeod, P., and Gobet, F. (2007a). Does Chess Need Intelligence? — A Study with Young Chess Players. *Intelligence*, 35(5):457.
- Bilalić, M., McLeod, P., and Gobet, F. (2007b). Personality Profiles of Young Chess Players. *Personality and Individual Differences*, 42(6):901–910.
- Bilen, E. and Matros, A. (2023). The Queen’s Gambit: Explaining the Superstar Effect using Evidence from Chess. *Journal of Economic Behavior & Organization*, 215(November):307–324.
- Blanch, A. and Llaveria, A. (2021). Ability and Non-Ability Traits in Chess Skill. *Personality and Individual Differences*, 179:110909.
- Brands, R. A. and Fernandez-Mateo, I. (2017). Leaning Out: How Negative Recruitment Experiences Shape Women’s Decisions to Compete for Executive Roles. *Administrative Science Quarterly*, 62(3):405–442.
- Bühren, C. and Frank, B. (2012). Chess Players’ Performance Beyond 64 Squares: A Case Study on the Limitations of Cognitive Abilities Transfer. *Talent Development and Excellence*, 4(2):157–169.

- Buser, T. (2016). How Does the Gender Difference in Willingness to Compete Evolve with Experience? Tinbergen Institute Discussion Paper 16-017/I, Tinbergen Institute, Amsterdam, The Netherlands.
- Buser, T. and Yuan, H. (2019). Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad. *American Economic Journal: Applied Economics*, 11(3):225–252.
- Chakraborty, P. and Serra, D. (2024). Gender and Leadership in Organisations: The Threat of Backlash. *The Economic Journal*, 134(660):1401–1430.
- Cotton, C., McIntyre, F., and Price, J. (2013). Gender Differences in Repeated Competition: Evidence from School Math Contests. *Journal of Economic Behavior & Organization*, 86:52–66.
- De Sousa, J. and Hollard, G. (2023). From Micro to Macro Gender Differences: Evidence from Field Tournaments. *Management Science*, 69(6):3358–3399.
- De Sousa, J. and Niederle, M. (2022). Trickle-Down Effects of Affirmative Action: A Case Study in France. NBER Working Paper 30367, National Bureau of Economic Research, Cambridge, MA, United States of America.
- Dilmaghani, M. (2021). A Matter of Time: Gender, Time Constraint, and Risk Taking among the Chess Elite. *Economics Letters*, 208:110085.
- Dilmaghani, M. (2022). Chess Girls Don’t Cry: Gender Composition of Games and Effort in Competitions among the Super-Elite. *Journal of Economic Psychology*, 89:102482.
- Doldor, E., Wyatt, M., and Silvester, J. (2019). Statesmen or Cheerleaders? Using Topic Modeling to Examine Gendered Messages in Narrative Developmental Feedback for Leaders. *The Leadership Quarterly*, 30(5):101308.
- Dreber, A., Gerdes, C., and Gränsmark, P. (2013a). Beauty Queens and Battling Knights: Risk Taking and Attractiveness in Chess. *Journal of Economic Behavior & Organization*, 90:1–18.
- Dreber, A., Gerdes, C., Gränsmark, P., and Little, A. (2013b). Facial Masculinity Predicts Risk and Time Preferences in Expert Chess Players. *Applied Economics Letters*, 20(16):1477–1480.
- Dupas, P., Sasser Modestino, A., Niederle, M., and Wolfers, J. (2021). Gender and the Dynamics of Economics Seminars. NBER Working Paper 28494, National Bureau of Economic Research, Cambridge, MA, U.S.A.
- Eberhardt, M., Facchini, G., and Rueda, V. (2023). Gender Differences in Reference Letters: Evidence from the Economics Job Market. *The Economic Journal*, 133(655):2676–2708.
- Elo, A. (1978). *The Rating of Chess Players Past & Present*. Arco Publishing, New York City, NY, United States of America, 2 edition.
- Ertac, S. and Szetes, B. (2011). The Effect of Information on Gender Differences in Competitiveness: Experimental Evidence. Working Paper 1104, Koc University-Tüsiad Economic Research Forum, Istanbul, Turkey.
- FIDE (2021). FIDE Title Regulations Effective from 1 January 2022. Technical report, International Chess Federation (FIDE), Lausanne, Switzerland.
- Flabbi, L., Macis, M., Moro, A., and Schivardi, F. (2019). Do Female Executives Make a Difference? The Impact of Female Leadership on Gender Gaps and Firm Performance. *The Economic Journal*, 129(August):2390–2423.

- Gerdes, C. and Gränsmark, P. (2010). Strategic Behavior across Gender: A Comparison of Female and Male Expert Chess Players. *Labour Economics*, 17(5):766–775.
- Gill, D. and Prowse, V. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *American Economic Review*, 102(1):469–503.
- Gill, D. and Prowse, V. (2014). Gender Differences and Dynamics in Competition: The Role of Luck. *Quantitative Economics*, 5(2):351–376.
- Glickman, M. E. (1995). A Comprehensive Guide to Chess Ratings. *American Chess Journal*, 3(1):59–102.
- Glickman, M. E. (1999). Parameter Estimation in Large Dynamic Paired Comparison Experiments. *Applied Statistics*, 48(3):377–394.
- Glickman, M. E. (2022). Example of the Glicko-2 System. Technical report, Boston University, Boston, MA, United States of America.
- Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.
- Gneezy, U. and Rustichini, A. (2004). Gender and Competition at a Young Age. *American Economic Review*, 94(2):377–381.
- Gneezy, U., Saccardo, S., and Pietrasz, A. (2017). On the Size of the Gender Difference in Competitiveness. *Management Science*, 64(4):1541–1554.
- Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *The American Economic Review*, 104(4):1091–1119.
- González-Díaz, J. and Palacios-Huerta, I. (2016). Cognitive Performance in Competitive Environments: Evidence from a Natural Experiment. *Journal of Public Economics*, 139:40–52.
- González-Díaz, J., Palacios-Huerta, I., and Abuín, J. M. (2024). Personal Bests and Gender. *The Review of Economics and Statistics*, 106(2):409–422.
- Grabner, R. H., Stern, E., and Neubauer, A. (2007). Individual Differences in Chess Expertise: A Psychometric Investigation. *Acta Psychologica*, 124(3):398–340.
- Haenni, S. (2019). Ever Tried. Ever Failed. No Matter? On the Demotivational Effect of Losing in Repeated Competitions. *Games and Economic Behavior*, 115:346–362.
- Kang, L., Lei, Z., Song, Y., and Zhang, P. (2024). Gender Differences in Reactions to Failure in High-Stakes Competition: Evidence from the National College Entrance Exam Retakes. *Journal of Political Economy Microeconomics*, 2(2).
- Koizumi, H. (2024). How Much of Merit is Due to Luck? Evidence on the Butterfly Effect of Luck. Discussion Paper 24035, The Research Institute of Economy, Trade and Industry, Tokyo, Japan.
- Künn, S., Palacios, J., and Pestel, N. (2023). Indoor Air Quality and Strategic Decision Making. *Management Science*, 69(9):5354–5377.
- Künn, S., Seel, C., and Zegners, D. (2022). Cognitive Performance in Remote Work - Evidence from Professional Chess. *The Economic Journal*, 132(643):1218–1232.
- Levitt, S. D., List, J., and Sadoff, S. E. (2011). Checkmate: Exploring Backward Induction among Chess Players. *American Economic Review*, 101(2):975–990.

- Maass, A., D’Ettola, C., and Cadinu, M. (2008). Checkmate? The Role of Gender Stereotypes in the Ultimate Intellectual Sport. *European Journal of Social Psychology*, 38:231–245.
- Markowsky, E. and Beblo, M. (2022). When Do We Observe a Gender Gap in Competition Entry? A Meta-Analysis of the Experimental Literature. *Journal of Economic Behavior & Organization*, 198(2022):139–163.
- Mueller, C. (2023). Reacting to Early Failure in University: Evidence from a Regression Discontinuity Design. Beiträge zur Jahrestagung des Vereins für Socialpolitik 2023: Growth and the ”sociale Frage”, ZBW - Leibniz Information Centre for Economics,, Kiel, Hamburg, Germany.
- Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3):1067–1101.
- Niederle, M., Vesterlund, L., and Segal, C. (2013). Gender Differences in Competition. *How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness*, 59(1):1–16.
- Niederle, M. and Yestrumskas, A. (2008). Gender Differences in Seeking Challenges: The Role of Institutions. NBER Working Paper 13922, National Bureau of Economic Research.
- Niessen-Ruenzi, A. and Ruenzi, S. (2019). Sex Matters: Gender Bias in the Mutual Fund Industry. *Management Science*, 65(7):3001–3025.
- Palacios-Huerta, I. (forthcoming). The Beautiful Dataset. *Journal of Economic Literature*.
- Palacios-Huerta, I. and Volij, O. (2009). Field centipedes. *American Economic Review*, 99(4):1619–35.
- Piasenti, S., Valente, M., van Veldhuizen, R., and Pfeifer, G. (2023). Does Unfairness Hurt Women? The Effects of Losing Unfair Competitions. IZA Discussion Paper 16324, IZA- Institute of Labor Economics, Bonn, Germany.
- Sarsons, H., Gërkhani, K., Reuben, E., and Schram, A. (2021). Gender Differences in Recognition for Group Work. *Journal of Political Economy*, 129(1):101–147.
- Schröder-Hansen, K. and Hansen, A. (2023). Performance Management Trends: Reflections on the Redesigns Big Companies Have Been Doing Lately. *International Journal of Productivity and Performance Management*, 72(5):1201–1220.
- Shastry, G. K., Shurchkov, O., and Xia, L. X. (2020). Luck or Skill: How Women and Men React to Noisy Feedback. *Journal of Behavioral and Experimental Economics*, 88:101592.
- Shurchkov, O. (2012). Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints. *Journal of the European Economic Association*, 10(5):1189–1213.
- Smerdon, D. (2022). The Effect of Masks on Cognitive Performance. *PNAS*, 119(49):e2206528119.
- Smerdon, D., Hu, H., McLennan, A., von Hippel, W., and Albrecht, S. (2020). Female Chess Players Show Typical Stereotype-Threat effects: Commentary on Stafford (2018). *Psychological Science*, 31(6):756–759.
- Strittmatter, A., Sunde, U., and Zegners, D. (2020). Life Cycle Patterns of Cognitive Performance over the Long Run. *PNAS*, 117(44):27255–27261.
- Unterrainer, J. M., Kaller, C. P., Rahm, B., and Halsband, U. (2006). Planning Abilities and Chess: A Comparison of Chess and Non-Chess Players on the Tower of London Task. *British Journal of Psychology*, 97(3):299–311.

- Vollstädt-Klein, S., Grimm, O., Kirsch, P., and Bilalić, M. (2010). Personality of Elite Male and Female Chess Players and its Relation to Chess Skill. *Learning and Individual Differences*, 20(5):517–521.
- Wasserman, M. (2023). Gender Differences in Politician Persistence. *The Review of Economics and Statistics*, 105(2):275–291.
- Waters, A. J., Gobet, F., and Leyden, G. (2002). Visuospatial Abilities of Chess Players. *British Journal of Psychology*, 93(4):557–565.
- Wozniak, D. (2012). Gender Differences in a Market with Relative Performance Feedback: Professional Tennis Players. *Journal of Economic Behavior & Organization*, 83(1):158–171.
- Wozniak, D., Harbaugh, W. T., and Mayr, U. (2016). The Effect of Feedback on Gender Differences in Competitive Choices. Working Paper.

A Data Appendix

A.1 Institutional Setting

A.1.1 Timeline Nationality Participation

On February 27, 2022, Chess.com released a statement condemning Russia’s invasion of Ukraine and, as part of its response, removed the Russian and Belarusian flags from player profiles, replacing them with a gray square. Subsequently, on April 23, 2022, the Russian government agency Roskomnadzor (the Federal Service for Supervision of Communications, Information Technology, and Mass Media) banned access to Chess.com in Russia. The ban was targeted at two specific pages on Chess.com: “On the Invasion of Ukraine,” which outlined the platform’s policies regarding the war, and “Ukrainian Chess Players in Times of War,” which reported on the experiences of Ukrainian players in the early days of the conflict. Because Chess.com operates on secure HTTPS webpages, Roskomnadzor could not block individual pages and therefore banned the site in its entirety. The ban can be circumvented by app use or VPN services.

The observation period therefore ends on April 17, 2022, the last tournament before the ban to ensure all players face similar barriers to entering the tournament. Figure A.1 plots the share of Russian players in the months following the Russian ban of the platform. While there was an immediate decrease following the ban, numbers have slowly recovered but fail to reach their old levels.

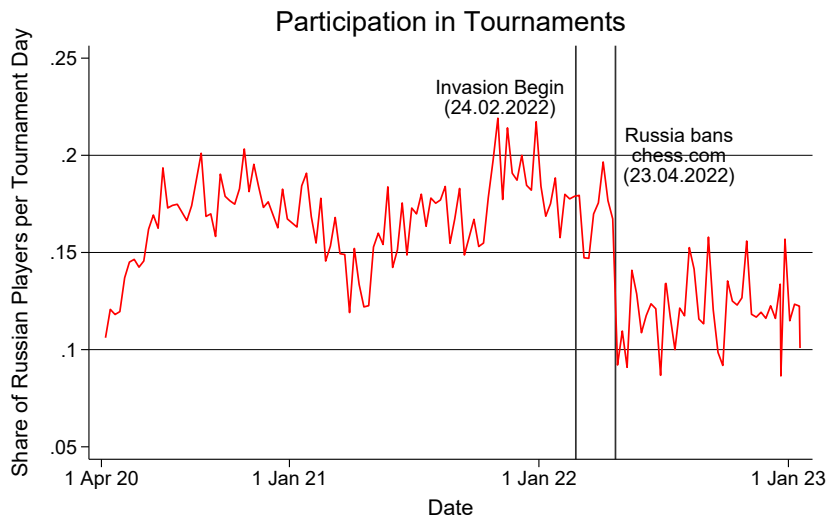


Figure A.1: Share of Russian Players

Note: Share of Russian players in each tournament, defined as the number of participants registered with the Russian Chess Federation divided by total tournament participants.

Since Russian female players make up 19% of all identified female players, extending the timeline beyond April 17, 2022 would no longer be comparable. Non-Russian players may change their competition decisions conditional on (stereotypically strong) Russian players no longer being in the running for the prize money.

A.1.2 Sonneborn-Berger Tiebreakers

The tournament is won by the player that achieves the most points. Should multiple players be tied for a place, tiebreakers are used to determine the final placements. The Titled Tuesday series uses the Sonneborn-Berger tiebreaker method. The basic idea behind the Sonneborn-Berger

tiebreaks is to reward players for defeating strong opponents. For each opponent that a player defeats, the player earns a certain number of tiebreak points, based on the opponent’s performance in the tournament. Specifically, the tiebreak points awarded to a player for a win over an opponent are equal to the opponent’s score. If the player draws with an opponent, they each receive half of the opponent’s score.

Once all tiebreak points have been calculated, a player’s total Sonneborn-Berger score is the sum of all their tiebreak points. The player with the highest Sonneborn-Berger score is then ranked higher in the tournament standings, as illustrated in Table A.1. If players still have the same Sonneborn-Berger score the prize is shared.

Table A.1: Example Tournament Score and Sonneborn-Berger Tiebreaks

Player	Points	Opponent	Total Points (Opponent)	Total Points (Player)	Sonneborn-Berger	Place
	0	B	2			
A	1	C	1.5	2	2	2
	1	D	0.5			
B	1	A	2			
	0.5	C	1.5	2	3	1
	0.5	D	0.5			
C	0	A	2			
	0.5	B	2	1.5	1.5	3
	1	D	0.5			
D	0	A	2			
	0.5	B	2	0.5	0.5	4
	0	C	1.5			

Note: The table presents the results of a fictitious four-person tournament and how to calculate the Sonneborn-Berger tiebreak scores.

A.1.3 Rating System

The rating system in chess evaluates a player’s skill relative to others. The most widely used rating system is the [Elo \(1978\)](#) rating system and subsequent refinements ([Glickman, 1995, 1999, 2022](#)). Each player receives a numerical rating with higher ratings indicating stronger players. Variations of this system are used in other sports, for example to determine tennis rankings.

Game outcomes (win, loss, or draw) affects player ratings. If a lower-rated player defeats a higher-rated player, they earn more rating points than if they were to defeat a player with a similar rating. Conversely, if a higher-rated player defeats a lower-rated player, they earn fewer rating points. Draws typically result in minor adjustments favoring the lower-rated player. The expected outcome of a match depends on the players’ rating difference: higher-rated players are more likely to win, and rating updates reflect the difference between actual and expected results. Over time, a player’s rating becomes a reliable measure of skill within the chess community.

The rating difference predicts the win probability of a match ([Elo, 1978](#)), as Equation [A.1](#) shows:

$$P(\text{Win})_{elo_{P1}, elo_{P2}} = \frac{1}{1 + 10^{\frac{elo_{P1} - elo_{P2}}{400}}} \quad (\text{A.1})$$

Players with equal ratings have a 50% chance of winning, while larger rating gaps correspond to higher expected win probabilities (e.g., a 300-point advantage implies an 85% chance of winning for the higher-rated player). For the dataset, each player’s Elo rating and the rating difference to their opponent are recorded to identify the presumed favorite in each match.

On Chess.com, Elo ratings measure skill relative to other platform users. Even in this elite field, I still record Elo differences of several 100 points. A player’s Elo rating is predictive of performance, as final scores closely mirror the Elo distribution, shown in Figure [A.2](#).

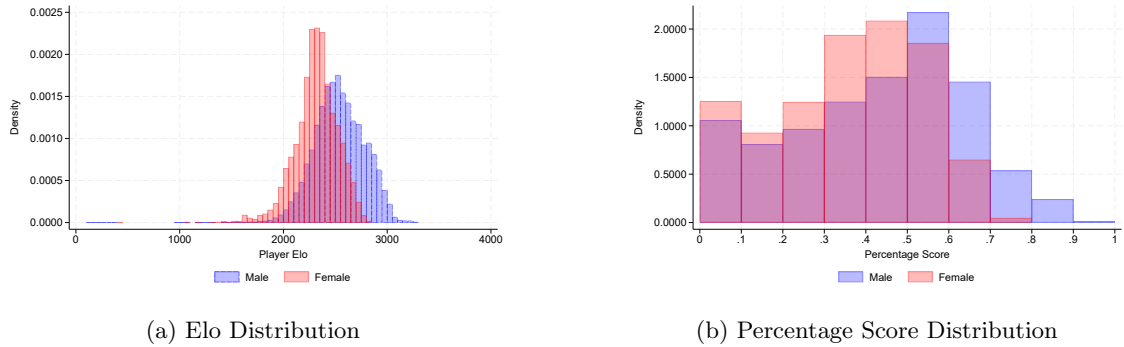


Figure A.2: Final Outcomes and Performance Measure

Note: Descriptive overview of skill and results distributions in the sample. Panel A ([A.2a](#)) shows a histogram of average within-player Elo ratings for 3,522 identified players, by gender, averaged across all games over a two-year period. Panel B ([A.2](#)) shows a histogram of percentage scores for all identified players, by gender, reported for every tournament played over the same period.

A.1.4 Player's Information Set



Figure A.3: Player Perspective Titled Tuesday Stream

Note: Screenshot from chess streamer Hikaru Nakamura's Titled Tuesday livestream on August 4, 2020. The right half of the image shows the in-game interface available to players, including the chessboard, clocks, titles, ratings, usernames, and profile pictures.

A.1.5 Chess Titles

In chess, titles serve as recognition of a player’s skill and achievement, most commonly awarded by the international chess federation FIDE. National chess federations may confer titles like National Master (NM), each with its own set of requirements and significance. These titles signify a player’s proficiency to chess community. The following lists describe the types of titles available and the requirements to obtain one, for more details see [FIDE \(2021\)](#).

To earn a title, players must fulfill specific requirements (“norms”). A norm is typically a rating requirement or a outstanding tournament performance against strong opponents. Norms serve as milestones on the path to earning a title and require a sustained high-level performance. Lower tier titles can be obtained by either achieving specific norms and reaching a minimum rating or by crossing a specific rating threshold.

1. Grandmaster (GM): The highest title bestowed by FIDE, indicating exceptional skill. To attain GM status, a player typically needs to achieve three GM norms, which are outstanding performances in international tournaments, and obtain a FIDE rating of 2500 or above.
2. International Master (IM): A prestigious title below Grandmaster. To become an IM, a player must achieve IM norms and obtain a FIDE rating of 2400 or above.
3. FIDE Master (FM): A significant title below International Master. To reach FM status, a player must achieve a FIDE rating of 2300 or above in any single FIDE rating list.
4. Candidate Master (CM): An entry-level title for serious tournament players. To be awarded the CM title immediately, a player must achieve a FIDE rating of 2200 or above in any single FIDE rating list.

Women’s titles in chess are similar to their counterparts in open chess, but exclusive to female players. Provided they fulfilled the norms, women are free to carry the open titles as well.

1. Woman Grandmaster (WGM): Equivalent to the Grandmaster title in open chess but specific to women. To become a WGM, a player typically needs to achieve three WGM norms (outstanding performances in international tournaments) and obtain a FIDE rating of 2300 or above. The requirements for WGM titles are slightly lower than those for GM titles.
2. Woman International Master (WIM): Similar to the International Master title in open chess but specific to women. To attain the WIM title, a player must achieve WIM norms and obtain a FIDE rating of 2200 or above.
3. Woman FIDE Master (WFM): Equivalent to the FIDE Master title in open chess but specific to women. Requirements for the WFM title are that a player reaches a FIDE rating of 2100 or above in any single FIDE rating list.
4. Woman Candidate Master (WCM): An entry-level title for female tournament players, similar to the Candidate Master title in open chess. To be awarded the WCM title, a player must achieve a FIDE rating of 2000 or above in any single FIDE rating list.

The Titled Tuesday series takes place online which limits visual cues about the opponent. Players only have information on the profile picture, rating, and title of the opponent. If a player is paired with an opponent with a female title whose acronyms’ are prefixed with a “W”, they know that they are playing against a woman. Not every woman chess player holds a female-specific title, but the majority does (85%, see [Table A.2](#)).

Table A.2: Title Distribution for Identified Players

	<u>Male</u>	<u>Female</u>
Grandmaster (GM)	0.205	0.035
International Master (IM)	0.226	0.082
FIDE Master (FM)	0.335	0.016
Candidate Master (CM)	0.107	0.002
National Master (NM)	0.127	0.019
Woman Grandmaster (WGM)		0.158
Woman International Master (WIM)		0.239
Woman FIDE Master (WFM)		0.331
Woman Candidate Master (WCM)		0.119
Observations	3008	514

Note: The table presents an overview of the title distribution for identified players by gender.

A.2 Variables

A.2.1 Centipawn Loss

Modern computers far exceed human chess-playing capabilities with the help of chess programs called “engines”. For this analysis the open-source chess engine Stockfish (Version 14.2) was used at depth 20, the strongest chess computer available in 2022. Unlike humans, computers have no intuition for good moves, but excel at calculation which is at the center of chess. To replace intuition, most engines are equipped with the rules of chess, as well as some guidelines regarding the pieces value.¹¹ These rules help an engine decide whether a move/possible move sequence is promising. Giving up a knight (3 points of material) for an opponent’s rook (5 points of material) likely secures an advantage and should be explored further. A move which loses a queen without anything in return costs 9 points of material (and likely the game) and should not be played.

For a given chess position, the computer provides an evaluation of the position. The computer will go through various possible moves, quickly discarding non-promising moves to save on computational power. For each move it will then iterate forward possible opponent’s responses, their possible response to that response, etc. The position evaluation therefore does not only include the current board position, but also the computer’s prediction on how this position will evolve under future optimal play. These processes require a lot of processing power and memory. To limit the amount of moves a computer can look forward the users have the option to limit the “search depth”. The depth for this paper is limited to 20 half-moves ahead to ensure that the calculation time is feasible.

While the computer’s assessment process does not lend itself to human interpretation, once the engine has finished its assessment it returns three measures that are understandable to chess players. Firstly, the machine will return a “score”. The score is a numeric assessment of which side is currently winning expressed in the unit of centipawns ($\frac{1}{100}$ of a pawn). A centipawn score of +100 signifies that White’s position is better to Black’s and the (positional) advantage is equivalent to White having an additional pawn on the board. As previously mentioned: the score that the engine returns already encompasses the engine’s assessment of how the game will evolve under best play from the assessed position onward.

There is a difference between the score and the material advantage held by the players. The material advantage is very inflexible and merely adds up the values of the different figures on the board. The score is more flexible, while it takes the worth of pieces into account (e.g. a queen is worth more than a knight) it also takes piece placement into account. If both players have equal material, but White has their pieces placed on better or more important squares, then the score will return an advantage for White.

However, using centipawn loss can be very punishing. For example, assume that White has a completely winning position and can force checkmate in 4 moves under ideal play. Stockfish would consider White to have made an error had they played any other move than a move resulting in checkmate in 3. In reality, short of leaving a piece to be captured White’s position will still be completely winning regardless of which move they make. Using the win-draw-loss statistic would show that White still has the ability to force checkmate and thus a winning probability of 100%. Because certain mates are not returned in centipawns, the primary accuracy measure used, I award

¹¹Decades of chess theory offers the following rules-of-thumb. A pawn (the weakest capture piece) is normed at a value of one and the value of every other piece is assessed relatively to that pawn. Knights and bishops are each worth 3 pawns. Rooks are worth 5 pawns. Queens are worth 9 pawns.

a centipawn loss of 0 if the player and the computer both have a certain mate, independent of in how many moves the position can be achieved. If a player blunders resulting in a position that gives the opponent the opportunity to force checkmate, I award a centipawn loss of 1,000, equivalent to a 10 pawn disadvantage and the slightly more than losing a queen. The centipawn loss I use is in line with the punishment parameters machines use to decide on the next move.

The final measure that Stockfish returns is the “best move”. For a given position on the board, Stockfish will identify the best move for whichever player is set to play next. This “best move” can then be played and compared to what actually happened next in the game. By comparing the scores of Stockfish’s preferred move and what the players played I am able to construct an accuracy measure describing how much advantage players gave up on by not finding the best move. Because Stockfish is superior to all chess players, this is a neutral benchmark and can be compared across players.

An example from one of the games in the dataset:

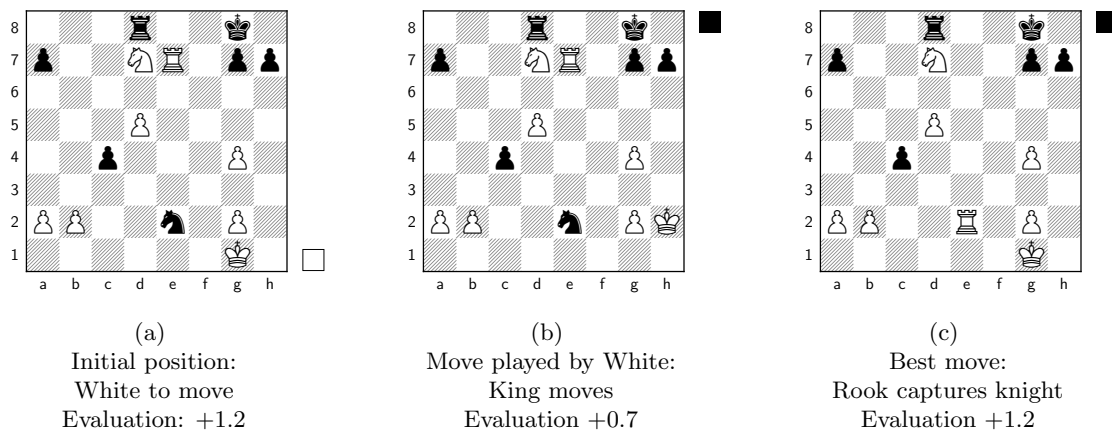


Figure A.4: Comparison Player vs. Computer Evaluation

Note: The figure offers a visual example of how the centipawn loss is calculated. Subfigure A.4a is the initial board position. Subfigure A.4b is the move that was actually played. Subfigure A.4c is the move that the computer identified as best. The centipawn loss associated with the move in Subfigure A.4b is the difference of the evaluation of Subfigure A.4c and Subfigure A.4c.

Let the initial position be Subfigure A.4a with White to move. The white king is in check because of the black knight, but the computer’s evaluation maintains that White has a slight advantage at +1.2 pawns. There are two options: the king must move out of check or the knight must be captured. In the game subfigure A.4b was played, the king was moved out of the way reducing the advantage White had to +0.7 pawns. The engine however prefers a move which involves capturing the knight with the white rook, according to the engine’s calculations this move would preserve the advantage that White had after Black’s move. Therefore, White gave up an advantage numerically equivalent to half a pawn.

A.2.2 Mistakes Distribution over Game

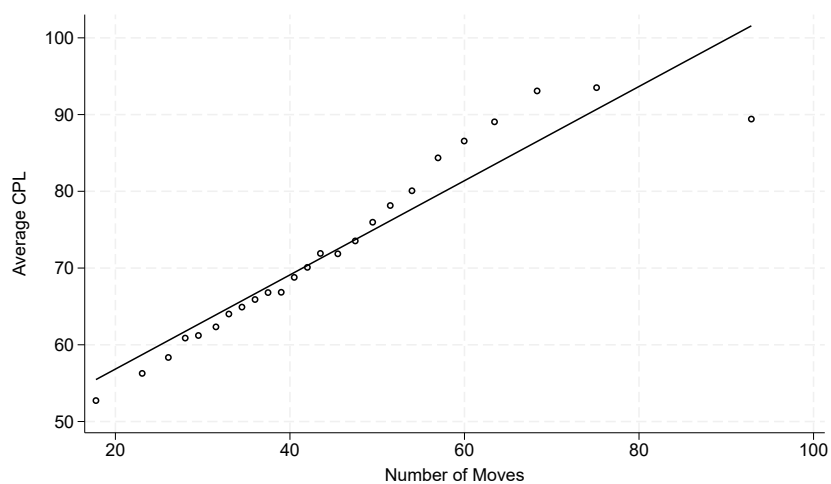


Figure A.5: Move Accuracy by Number of Moves in Game

Note: The figure presents a binscatter of 25 bins of the average centipawn loss by number of moves for the full analytical sample. Every point represents the within-quantile mean.

A.2.3 Centipawn Loss and Game Outcomes

I run the following regression on the points won by player i in a game played in round r of tournament t for the analytical sample derived in Section 4.1. The points follow the conventional chess scoring system, where a win corresponds to 1 point, a draw to half a point, and a loss results in 0 points.

$$points_{irt} = \alpha_i + \beta_1 \overline{CPLoss}_{irt} + \beta_2 \overline{CPLoss}_{irt}^2 + \Theta \mathbf{X}_{irt} + round_r + tournament_t + \varepsilon_{irt} \quad (\text{A.2})$$

Where \overline{CPLoss}_{irt} and its square is the average centipawn loss per move in a game played. The vector \mathbf{X}_{irt} controls for the player's elo rating, the rating difference to their opponent, opponent characteristics (title and profile picture content), and the color of the playing pieces. I use additional player, round, and tournament fixed effects.

Table A.3: Relationship of Centipawn Loss and Game Outcomes

	(1)	(2)	(3)
Capped Avg. CP Loss	-0.005*** (0.000)	-0.010*** (0.000)	
Capped Avg. CP Loss ²		0.00003*** (0.000)	
Elo Difference (in 100s)	0.092*** (0.001)	0.089*** (0.001)	0.101*** (0.001)
White Pieces	0.046*** (0.002)	0.045*** (0.002)	0.049*** (0.002)
Observations	312,822	312,822	312,822
Adjusted R ²	0.316	0.332	0.216

Note: The dependent variable is points achieved in a round (1 = win, 0.5 = draw, 0 = loss). Average centipawn loss per move measures the difference between the advantage of the computer's preferred move and the advantage actually obtained by the player; higher values indicate greater deviation from the computer's recommendation. All regressions control for the player's rating, the rating difference to the opponent, age, the opponent's title, and the gender signaled in the opponent's profile picture. Regressions include player, tournament, and round fixed effects. Standard errors, clustered at the individual level, are shown in parentheses. Significance levels: $p < 0.10$, $p < 0.05$, $p < 0.01$.

A.2.4 Centipawn Loss and Right-Censoring

The evaluation in centipawns is theoretically bounded to infinity until the computer sees a forced checkmate sequence in the number of half-moves it looks ahead (for this paper: 20 half-moves). As such, if a player makes a significant mistake from a winning position to a losing position without it resulting in a forced checkmate the centipawn loss with the computer can be very high. For example let the evaluation currently sit at 2,000cp and the player makes a significant error that results in an evaluation of $-2,000cp$ after their move. The centipawn loss of this individual move would be 4,000 and the player is now losing. A similar statement could be made for a move taking a 1,000cp evaluation to a $-1,000cp$ evaluation though this mistake only carries a magnitude of 2,000cp. The magnitude of these losses is similar in terms of their impact on winning chances, but have different centipawn losses associated with them. By restricting the right tail end to a maximum centipawn loss to 1000, the variance of the outcome variable is restricted and the analysis becomes less sensitive to outliers.

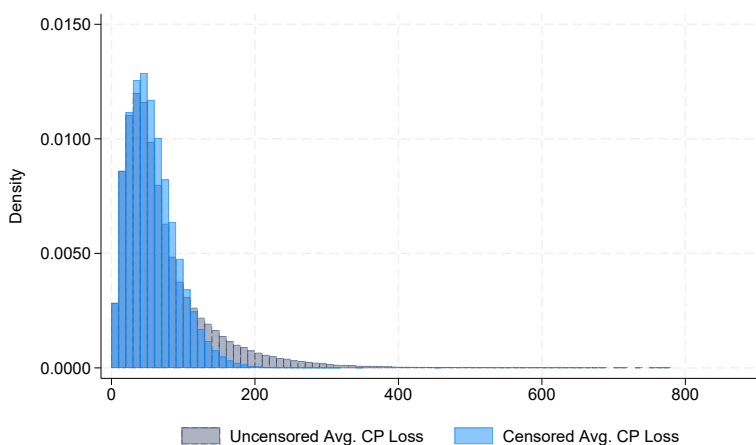


Figure A.6: Comparison Raw vs. Censored Centipawn Loss

Note: Histogram of average centipawn loss for the analytical sample. Games include only those with more than 10 moves played by identified players after the third round. Centipawn loss is censored at 1,000cp to limit the magnitude of move-level mistakes (blue); uncensored values are shown in gray.

To address the possibility that my results are driven by games featuring these extreme losses. I modify the average centipawn loss in a dependent variable to limit the maximum centipawn loss of a single move at 1000. This reduces the values at the right tail end, as evidenced in Table A.4 and Figure A.6.

Table A.4: Summary Statistics - Game Level

	Male	Female	Full Sample					
	Mean	Mean	Mean	Median	95. Perc.	99. Perc.	SD	N
Average CPL	70.498	80.634	71.953	53.788	196.385	302.535	61.525	312892
Average Capped CPL	55.822	64.516	57.070	51.575	120.026	157.514	33.627	312892

Note: Descriptive statistics of average centipawn loss in a game for the analytical sample. The average capped centipawn loss right-censors the maximum centipawn loss at 1000. Sample consists of games played after the third round by identified players with more than 10 moves. Means reported by gender in the two leftmost columns.

A.2.5 Opening Choice and Game Outcomes

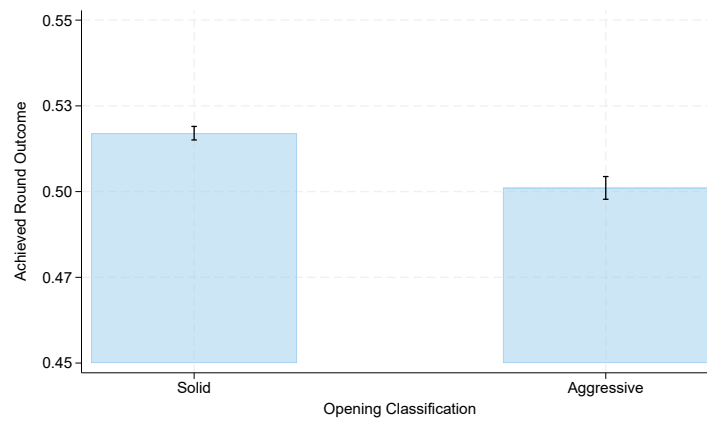


Figure A.7: Game Outcome and Opening Choice

Note: Average points won per round by opening type, based on the classification of Gerdes and Gränsmark (2010).

A.2.6 Percentage Score

Because there are only three possible point outcomes for each game of chess (0, 0.5, or 1) there is only a finite set of possible points that players end the tournament with. Table A.5 gives an overview of the possible percentage score a player can have within the Titled Tuesday series. Because the series changed the possible number of rounds twice, the percentage scores are adjusted accordingly.

Table A.5: Possible Percentage Scores

Points	Rounds in Tournament		
	9	10	11
0	0	0	0
0.5	0.06	0.05	0.05
1	0.11	0.1	0.09
1.5	0.17	0.15	0.14
2	0.22	0.2	0.18
2.5	0.28	0.25	0.23
3	0.33	0.3	0.27
3.5	0.39	0.35	0.32
4	0.44	0.4	0.36
4.5	0.50	0.45	0.41
5	0.56	0.5	0.45
5.5	0.61	0.55	0.50
6	0.67	0.6	0.55
6.5	0.72	0.65	0.59
7	0.78	0.7	0.64
7.5	0.83	0.75	0.68
8	0.89	0.8	0.73
8.5	0.94	0.85	0.77
9	1	0.9	0.82
9.5		0.95	0.86
10		1	0.91
10.5			0.95
11			1

Note: This table presents an overview of all possible percentage scores, given the total number of rounds in the tournament.

A.2.7 FIDE Performance Adjustment Factor

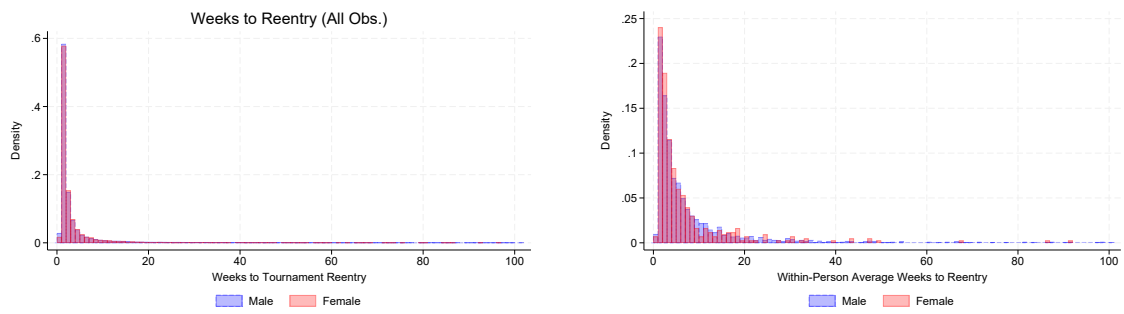
Table A.6: Rating Performance Adjustment Factor

Percentage Score	dp (Adjustment Factor)	Percentage Score	dp (Adjustment Factor)
1	800	0.49	-7
0.99	677	0.48	-14
0.98	589	0.47	-21
0.97	538	0.46	-29
0.96	501	0.45	-36
0.95	470	0.44	-43
0.94	444	0.43	-50
0.93	422	0.42	-57
0.92	401	0.41	-65
0.91	383	0.4	-72
0.9	366	0.39	-80
0.89	351	0.38	-87
0.88	336	0.37	-95
0.87	322	0.36	-102
0.86	309	0.35	-110
0.85	296	0.34	-117
0.84	284	0.33	-125
0.83	273	0.32	-133
0.82	262	0.31	-141
0.81	251	0.3	-149
0.8	240	0.29	-158
0.79	230	0.28	-166
0.78	220	0.27	-175
0.77	211	0.26	-184
0.76	202	0.25	-193
0.75	193	0.24	-202
0.74	184	0.23	-211
0.73	175	0.22	-220
0.72	166	0.21	-230
0.71	158	0.2	-240
0.7	149	0.19	-251
0.69	141	0.18	-262
0.68	133	0.17	-273
0.67	125	0.16	-284
0.66	117	0.15	-296
0.65	110	0.14	-309
0.64	102	0.13	-322
0.63	95	0.12	-336
0.62	87	0.11	-351
0.61	80	0.1	-366
0.6	72	0.09	-383
0.59	65	0.08	-401
0.58	57	0.07	-422
0.57	50	0.06	-444
0.56	43	0.05	-470
0.55	36	0.04	-501
0.54	29	0.03	-538
0.53	21	0.02	-589
0.52	14	0.01	-677
0.51	7	0	-800
0.5	0		

Note: The table provides an overview of the adjustment factor used to calculate FIDE's performance measure, as outlined in the [FIDE \(2021\)](#) title regulations.

The FIDE (tournament) performance rating is a special variation of the tournament performance rating formula. The standard tournament performance rating formula yields a rating of ∞ for perfect records and $-\infty$ if every game was lost. (FIDE, 2021) recognizes this, the +/- 800 adjustment factor is a (more or less arbitrary) notion choice for practical reasons. Having a defined adjustment enables players with perfect win records can use this performance to fulfill chess title norms. For a more detailed discussion and estimation strategies, I refer to Ismail (2023) for an overview.

A.2.8 Reentry Waiting Time Distribution



(a) Days until Reentry by Gender

(b) Days until Reentry by Individual Player

Figure A.8: Waiting Time Distribution

Note: Histograms of Titled Tuesday participation by gender. Panel A shows reentry choices recorded between April 14, 2020 (one week after the first tournament), and April 24, 2022. Panel B shows the distribution of average within-player waiting times, based on 2,580 players with at least two tournaments on record.

A.2.9 Rating Differences

As the tournament progresses, the pairing system increasingly matches players with opponents of similar skill, as determined by their Elo ratings. Consequently, the differences in Elo between opponents tend to decrease over successive rounds, as shown in Figure A.9.

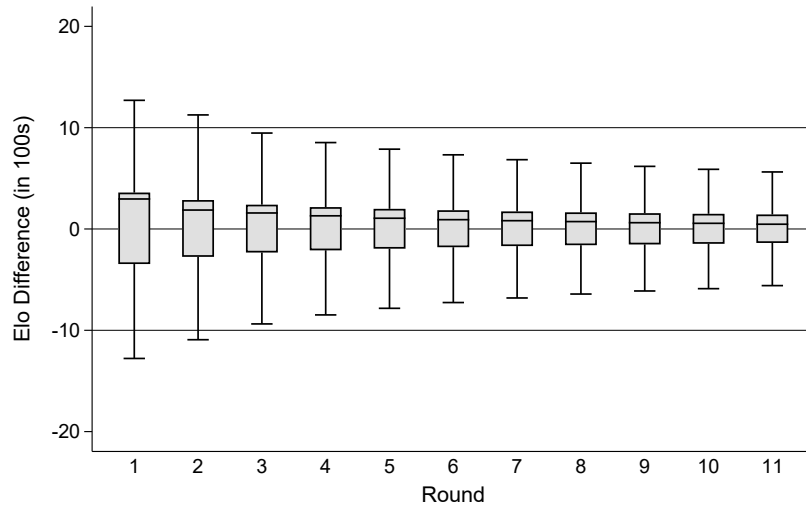
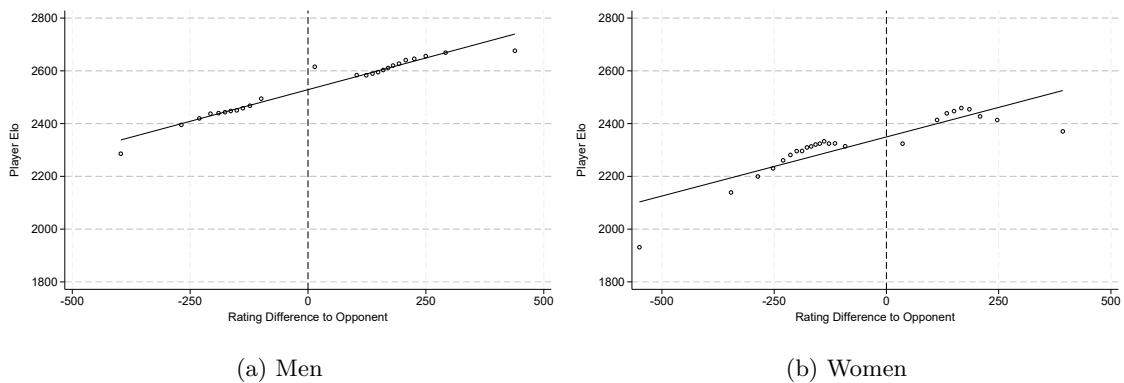


Figure A.9: Elo Difference By Round

Note: Figure shows a box plot showing the 95% confidence interval of the Elo difference to the opponent by round. The horizontal line within each box represents the median. Outliers are not displayed.



(a) Men

(b) Women

Figure A.10: Elo Difference to Opponent by Rating

Note: The figure presents a binscatter of a player's rating and the rating difference to their opponent separately by gender. Subfigure A.10a presents a binscatter in 25 bins of identified female players' ratings by elo difference to their opponent for the (female) analytical sample. Subfigure A.10b is a binscatter in 25 bins of identified female players' ratings by elo difference to their opponent for the (female) analytical sample. In both subfigures, each point represents the within-quantile mean.

B Robustness Checks

B.1 Centipawn Loss

B.1.1 Censored Centipawn Loss

The main results use the raw centipawn loss values and men’s cognitive performance decreases after a loss while women’s increases. As discussed in Appendix A.2.4, this can be a very punishing metric in the endgame. Using the censored centipawn loss values I find that the effect is weaker, but the result of men worsening and women improving remains.

Table B.1: Effect of Previous Rounds on Average Centipawn Loss

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.301** (0.138)	0.435*** (0.148)	1.114*** (0.200)	1.203*** (0.207)	-0.429*** (0.146)	-0.373** (0.160)
Lost R-2	0.602*** (0.138)	0.628*** (0.148)	1.312*** (0.213)	1.364*** (0.226)	-0.140 (0.154)	-0.222 (0.167)
Lost R-3	0.416*** (0.139)	0.475*** (0.149)	0.904*** (0.203)	1.071*** (0.213)	-0.128 (0.147)	-0.199 (0.161)
Lost R-1 × Female		-0.872** (0.357)		-0.752 (0.670)		-0.319 (0.371)
Lost R-2 × Female		-0.144 (0.360)		-0.383 (0.656)		0.459 (0.383)
Lost R-3 × Female		-0.374 (0.379)		-1.634** (0.708)		0.375 (0.375)
Observations	312,822	312,822	312,822	312,822	312,822	312,822
R ²	0.098	0.098	0.098	0.098	0.098	0.098

Note: Dependent variable is the censored average centipawn loss per move. A positive value implies deviation from the computer-preferred move resulting in an avoidable disadvantage. All regressions control for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player, tournament and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.1.2 Opponent Gender

The male-dominated chess environment and the stereotype of women being bad at chess may give rise to a differential response to a loss depending on the opponent type. Because the title is prominently displayed during a match, players facing a “W” title receive a unmistakable signal about their opponent’s gender and circumvents concerns about differing information content and image quality in profile pictures. This analysis has a drawback: approximately 15% of the players are female, of which a subset holds a women-specific title. In the analytical sample, 10.4% of games feature a opponent with a “W” title. On average, 6% of the games in the sample feature a loss against a known female opponent.

Table B.2: Effect of Opponent Gender

	Any Loss		vs. Female Title		vs. Open Title	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.242 (0.258)	0.476* (0.277)	1.012 (0.644)	0.624 (0.745)	-0.001 (0.251)	0.086 (0.253)
Lost R-2	0.948*** (0.249)	0.970*** (0.269)	1.134 (0.696)	0.900 (0.786)	0.716*** (0.246)	0.787*** (0.246)
Lost R-3	0.451* (0.248)	0.500* (0.265)	0.501 (0.720)	0.163 (0.822)	0.313 (0.245)	0.362 (0.247)
Lost R-1 × Female		-1.533** (0.655)		1.579 (1.451)		2.218* (1.260)
Lost R-2 × Female		-0.118 (0.664)		0.947 (1.672)		2.285 (1.479)
Lost R-3 × Female		-0.305 (0.711)		1.416 (1.693)		1.877 (1.492)
Observations	312,818	312,818	312,818	312,818	312,818	312,818
R ²	0.048	0.048	0.048	0.048	0.048	0.048

Note: Dependent variable is the average centipawn loss in a game. Sample consists of games played by identified players, excluding the first three rounds. All regressions include additional controls for the player’s own rating, the rating difference to their opponent their age, their opponent’s title and the gender displayed in the opponent’s profile picture. Opponent gender is defined as female if the opponent carries a female chess title, indicated with the title prefix “W”. All regressions include player, round and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Columns (1) and (2) in Table B.2 are the main results reported in Table 3 and are robust to further specifications. Men’s cognitive performance declines in the immediate aftermath of a loss whereas women’s improves. Columns (3) and (4) consider losses was against a player with a female title and find no detectable effect. Columns (5) and (6) consider losses against open titles, which can be held by men or women. For women, I find suggestive evidence that losing against an open title is associated with a decrease in move quality close to an upset loss.

Overall, the gender of the opponent does appears to matter more for the effect of losses. These results are surprising in the stereotypically male chess context. As the tournament takes place online it could be that players do not take notice of their opponent’s title or remember their opponent’s gender in the next match but only their rating. This would be different to an over-the-board setting where players must physically face each other while playing (Backus et al., 2023; Dreber et al., 2013a,b; Gerdes and Gränsmark, 2010).

B.1.3 Reaction to Winning

To check whether the between-round results are the result of behavioral reactions to losses the analysis is repeated for previous wins. The underlying regression equation follows the structure of Equation 11. The indicator variable for losses is exchanged for a variable which tracks past wins.

The main results for cognitive performance suggested that cognitive performance declines after a loss. Players who lost increase the level of inaccurate play by an average of half a centipawn per move for men. Women improved the quality of their play by approximately the same margin. Shifting the analysis towards winning reveals no such pattern. There is evidence that following a win as the underdog (i.e. facing an opponent with higher elo) results in a significant boost in cognitive performance in the next round for both men and women.

Table B.3: Effects of Previous Wins on Cognitive Performance

	Any Win		Underdog Victory	
	(1)	(2)	(3)	(4)
Won R-1	0.040 (0.242)	-0.050 (0.259)	-1.546*** (0.351)	-1.626*** (0.392)
Won R-2	-0.576** (0.241)	-0.539** (0.261)	-1.188*** (0.344)	-1.214*** (0.386)
Won R-3	-0.058 (0.239)	-0.064 (0.258)	-0.624* (0.359)	-0.589 (0.404)
Won R-1 × Female		0.644 (0.616)		0.421 (0.831)
Won R-2 × Female		-0.269 (0.627)		0.141 (0.823)
Won R-3 × Female		0.040 (0.645)		-0.166 (0.864)
Observations	312,818	312,818	312,818	312,818
R ²	0.048	0.048	0.048	0.048

Note: Dependent variable the average centipawn loss per move. A positive value implies deviation from the computer-preferred move resulting in an avoidable disadvantage. All regressions control for the player's own rating, the rating difference to their opponent, their age, their opponent's title and the gender signaled in the opponent's profile picture. Regressions without player fixed effects include time-invariant controls for the player's chess title, chess federation and gender. All regressions use tournament and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.1.4 Tournament-Round Fixed Effects

The main specification in Equation 11 uses separate fixed effects for round and tournament fixed effects. Such a specification may be more appropriate if there is a concern that the sample composition in a specific tournament influences the players differently across specific rounds. Examples would include the presence of a “superstar” in the tournament pool (Bilen and Matros, 2023). Given the size of the sample, it is possible to use a more restrictive variation by using $\text{tournament}_t \times \text{round}_r$ fixed effects with 774 combinations. Overall, selecting this fixed effect does not fundamentally change any of the conclusions in the paper.

Table B.4: Effect of Previous Rounds on Average Centipawn Loss with Tournament-Round FE

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.239 (0.258)	0.485* (0.278)	1.094*** (0.364)	1.274*** (0.381)	-0.559** (0.266)	-0.440 (0.293)
Lost R-2	0.948*** (0.249)	0.987*** (0.269)	1.952*** (0.405)	2.158*** (0.429)	-0.025 (0.275)	-0.211 (0.301)
Lost R-3	0.444* (0.248)	0.500* (0.265)	0.709* (0.377)	0.918** (0.393)	-0.062 (0.268)	-0.186 (0.296)
Lost R-1 \times Female		-1.603** (0.652)		-1.533 (1.188)		-0.681 (0.662)
Lost R-2 \times Female		-0.228 (0.666)		-1.812 (1.261)		1.051 (0.691)
Lost R-3 \times Female		-0.351 (0.707)		-1.989 (1.343)		0.651 (0.675)
Observations	312,818	312,818	312,818	312,818	312,818	312,818
R ²	0.049	0.049	0.050	0.050	0.049	0.049

Note: Dependent variable the average centipawn loss per move. A positive value implies deviation from the computer-preferred move resulting in an avoidable disadvantage. All regressions control for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player fixed effects and tournament \times round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Opening Aggression

B.2.1 Alternative Opening Aggression

The opening risk metric follows [Gerdes and Gränsmark \(2010\)](#). Eight chess experts (three women, five men, Elo ratings 2000–2600) classified openings as “aggressive” or “solid” from White’s and Black’s perspective. Openings are defined by move sequences, usually up to 15 moves (30 half-moves), and coded in the Encyclopedia of Chess Openings (ECO). Out of 499 ECO codes, the panel assigned a classification when at least six of eight experts agreed; otherwise, the opening remained ambiguous.

Of the 499 ECO codes, 26 could not be classified for either color. White has 125 ambiguous openings; Black, 150. The main analysis excludes these openings. To check robustness, I repeat the exercise using the full set of games. In this specification, openings fall into three categories: aggressive, solid, or ambiguous. The aggression dummy equals one if an aggressive opening is used and zero otherwise, so a positive coefficient means players are more likely to choose an “outright” aggressive opening compared to any non-aggressive option.

Table B.5: Effects of Previous Losses on Modified Opening Aggression

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.005*** (0.002)	0.006*** (0.002)	0.008*** (0.002)	0.010*** (0.003)	0.001 (0.002)	0.001 (0.002)
Lost R-2	0.004** (0.002)	0.004** (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
Lost R-3	0.000 (0.001)	0.001 (0.002)	-0.000 (0.002)	0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Lost R-1 × Female		-0.008* (0.004)		-0.013* (0.007)		-0.002 (0.004)
Lost R-2 × Female		-0.001 (0.004)		0.003 (0.008)		0.001 (0.004)
Lost R-3 × Female		-0.002 (0.004)		-0.015** (0.007)		0.003 (0.004)
Observations	312,822	312,822	312,822	312,822	312,822	312,822
R ²	0.069	0.069	0.069	0.069	0.069	0.069

Note: Dependent variable is a binary variable whether the opening is “aggressive”. Based on a modification of the initial opening classification presented by [Gerdes and Gränsmark \(2010\)](#). If a opening is deemed ambiguous, it is coded as a 0 for not being widely considered as “aggressive”. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use tournament and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results of the modified opening analysis suggest that the results are robust. Losses are associated with an increase in subsequent opening aggression in men, with the effect being close to nonexistent for women. The significance levels are lower, from a 5% level in the main analysis. The pattern remains the same. The effect of an upset loss is especially pronounced, as before.

B.2.2 Opponent Gender

To check for differential effects based on opponent gender, I further examine how losses against players with a female/open title affect opening strategy in the next round. Table B.6 summarizes the results. Columns (1) and (2) report the baseline estimates from Table ???. I find no systematic evidence that opponent gender matters when considering the losses.

Table B.6: Effect of Opponent Gender

	Any Loss		vs. Female Title		vs. Open Title	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.009*** (0.002)	0.011*** (0.002)	0.005 (0.005)	0.004 (0.006)	0.008*** (0.002)	0.008*** (0.002)
Lost R-2	0.008*** (0.002)	0.009*** (0.002)	0.002 (0.006)	0.005 (0.006)	0.007*** (0.002)	0.007*** (0.002)
Lost R-3	0.002 (0.002)	0.002 (0.002)	0.011* (0.006)	0.012* (0.007)	-0.000 (0.002)	-0.000 (0.002)
Lost R-1 × Female		-0.013** (0.006)		0.002 (0.012)		0.011 (0.011)
Lost R-2 × Female		-0.007 (0.006)		-0.013 (0.012)		-0.003 (0.011)
Lost R-3 × Female		-0.003 (0.006)		-0.004 (0.013)		0.009 (0.011)
Observations	203,544	203,544	203,544	203,544	203,544	203,544
R ²	0.146	0.146	0.146	0.146	0.146	0.146

Note: Dependent variable is a binary variable whether the opening is “aggressive”. Sample consists of games played by identified players, excluding the first three rounds. All regressions include additional controls for the player’s own rating, the rating difference to their opponent their age, their opponent’s title and the gender displayed in the opponent’s profile picture. Opponent gender is defined as female if the opponent carries a female chess title, indicated with the title prefix “W”. All regressions include player, round and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2.3 Reaction to Winning

To check whether the between-round results are the result of behavioral reactions to losses the analysis is repeated for previous wins. The underlying regression equation follows the structure of Equation 11. The indicator variable for losses is exchanged for a variable which tracks past wins.

The main results indicated that men increase the likelihood of using an aggressive opening after a loss, while women’s propensity to use aggressive openings does not change. Examining wins instead of losses inverts this pattern. Men are 0.7pp less likely to use an aggressive opening following a win and there is nearly no effect for women. The main analysis was able to identify a strong effect for upset losses, resulting in a 2pp increase in the likelihood of an aggressive opening. I am unable to detect any effect of an underdog victory on opening behavior.

Table B.7: Effects of Victories on Opening Aggression

	Any Win		Underdog Victory	
	(1)	(2)	(3)	(4)
Won R-1	-0.006*** (0.002)	-0.007*** (0.002)	0.001 (0.003)	0.000 (0.003)
Won R-2	-0.005** (0.002)	-0.005** (0.002)	0.004 (0.003)	0.002 (0.003)
Won R-3	0.001 (0.002)	-0.001 (0.002)	-0.002 (0.003)	-0.004 (0.003)
Won R-1 × Female		0.008 (0.006)		0.003 (0.008)
Won R-2 × Female		0.004 (0.006)		0.007 (0.008)
Won R-3 × Female		0.007 (0.005)		0.011 (0.008)
Observations	203,544	203,544	203,544	203,544
R ²	0.146	0.146	0.146	0.146

Note: Dependent variable is a binary variable indicating whether the opening is deemed “aggressive”. Based on a modification of the initial opening classification presented by [Gerdes and Gränsmark \(2010\)](#). If a opening is deemed ambiguous, it is coded as a 0 for not being widely considered as “aggressive”. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use tournament and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2.4 Tournament-Round Fixed Effects

The main specification in Equation 11 uses separate fixed effects for round and tournament fixed effects. Such a specification may be more appropriate if there is a concern that the sample composition in a specific tournament influences the players differently across specific rounds. Examples would include the presence of a “superstar” in the tournament pool (Bilen and Matros, 2023). Given the size of the sample, it is possible to use a more restrictive variation by using $\text{tournament}_t \times \text{round}_r$ fixed effects with 774 combinations. Overall, selecting this fixed effect does not fundamentally change any of the conclusions in the paper: Losing increases the likelihood of using an aggressive opening and the effect is particularly pronounced for upset losses.

Table B.8: Effects of Previous Losses on Opening Aggression with Tournament-Round FE

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.009*** (0.002)	0.011*** (0.002)	0.016*** (0.003)	0.018*** (0.004)	0.001 (0.002)	0.001 (0.002)
Lost R-2	0.008*** (0.002)	0.009*** (0.002)	0.004 (0.003)	0.004 (0.003)	0.005** (0.002)	0.005** (0.002)
Lost R-3	0.002 (0.002)	0.002 (0.002)	0.000 (0.003)	0.003 (0.003)	-0.001 (0.002)	-0.002 (0.002)
Lost R-1 \times Female		-0.013** (0.006)		-0.019** (0.010)		-0.004 (0.006)
Lost R-2 \times Female		-0.007 (0.006)		-0.001 (0.011)		-0.003 (0.006)
Lost R-3 \times Female		-0.003 (0.006)		-0.026** (0.010)		0.007 (0.006)
Observations	203,544	203,544	203,544	203,544	203,544	203,544
R ²	0.147	0.147	0.147	0.147	0.147	0.147

Note: Dependent variable is a binary variable indicating whether the opening is deemed “aggressive”. Based on a modification of the initial opening classification presented by Gerdes and Gränsmark (2010). If a opening is deemed ambiguous, it is coded as a 0 for not being widely considered as “aggressive”. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player fixed effects and tournament \times round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Quitting Behavior

B.3.1 Opponent Gender

Studying quitting behavior by game outcome and opponent gender, I find a consistent pattern with Table 5: the most recent result has the strongest effect. Notably, losing to a female-titled player (Column 2) has a smaller point estimate than losing to an open-titled player (Column 5). I cannot determine the differential effect of losing to a female title for women because of sample size limitations. Women quit less often overall, only 15% of the sample is female, and the test requires two women to be paired with variation in quitting behavior.

Table B.9: Effects of Opponent Gender on Quitting Behavior

	Any Loss		vs. Female Title		vs. Open Title	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.122*** (0.003)	0.129*** (0.004)	0.117*** (0.012)	0.117*** (0.012)	0.173*** (0.005)	0.175*** (0.006)
Lost R-2	0.020*** (0.001)	0.021*** (0.001)	0.017*** (0.006)	0.017*** (0.006)	0.022*** (0.002)	0.022*** (0.002)
Lost R-3	0.012*** (0.001)	0.011*** (0.001)	-0.007 (0.006)	-0.007 (0.006)	0.007*** (0.002)	0.006*** (0.002)
Lost R-1 × Female		-0.049*** (0.006)				-0.063*** (0.021)
Lost R-2 × Female		-0.004 (0.003)				0.003 (0.007)
Lost R-3 × Female		0.002 (0.003)				0.011 (0.008)
Observations	278,887	278,887	278,887	278,887	278,887	278,887
R ²	0.128	0.129	0.075	0.075	0.129	0.129

Note: Dependent variable is a binary variable indicating whether a player interrupted their tournament play. All regressions include controls for the player's own rating, the rating difference to their opponent, their age, their opponent's title and the gender signaled in the opponent's profile picture. All regressions use player, tournament, and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3.2 Reaction to Winning

To check whether the between-round results are the result of behavioral reactions to losses the analysis is repeated for previous wins. The underlying regression equation follows the structure of Equation 11. The indicator variable for losses is exchanged for a variable which tracks past wins.

The main results reported in Section 6 highlight that the chance to abandon the tournament is highest immediately after losing the game. To check that this behavior is driven through losses, Table B.10 repeats the analysis of quitting behavior for wins. As a sanity check, wins should either decrease the predicted probability of quitting or have no impact (given the low incidence). The results show that winning previous rounds reduces the probability of quitting the tournament, although for women the effect size is only about a third or half of men’s effect size.

Table B.10: Effects of Victories on Quitting Behavior

	Any Win		Underdog Victory	
	(1)	(2)	(3)	(4)
Won R-1	-0.109*** (0.003)	-0.115*** (0.003)	-0.057*** (0.002)	-0.061*** (0.002)
Won R-2	-0.021*** (0.001)	-0.022*** (0.001)	-0.031*** (0.002)	-0.031*** (0.002)
Won R-3	-0.014*** (0.001)	-0.014*** (0.001)	-0.017*** (0.001)	-0.016*** (0.002)
Won R-1 × Female		0.043*** (0.005)		0.021*** (0.004)
Won R-2 × Female		0.004 (0.003)		-0.001 (0.003)
Won R-3 × Female		0.0001 (0.003)		-0.0057* (0.003)
Observations	278,887	278,887	278,887	278,887
R ²	0.117	0.118	0.081	0.081

Note: Dependent variable is a binary variable indicating whether a player interrupted their tournament play. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player, tournament, and round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3.3 Tournament-Round Fixed Effects

The main specification in Equation 11 uses separate fixed effects for round and tournament fixed effects. Such a specification may be more appropriate if there is a concern that the sample composition in a specific tournament influences the players differently across specific rounds. Examples would include the presence of a “superstar” in the tournament pool (Bilen and Matros, 2023). Given the size of the sample, it is possible to use a more restrictive variation by using $\text{tournament}_t \times \text{round}_r$ fixed effects with 774 combinations. Overall, selecting this fixed effect does not fundamentally change any of the conclusions in the paper.

Table B.11: Effect of Previous Outcomes on Quitting Behavior

	Any Loss		Upset Loss		Expected Loss	
	(1)	(2)	(3)	(4)	(5)	(6)
Lost R-1	0.122*** (0.003)	0.129*** (0.004)	0.171*** (0.005)	0.176*** (0.006)	0.049*** (0.002)	0.051*** (0.002)
Lost R-2	0.020*** (0.001)	0.021*** (0.001)	0.024*** (0.002)	0.023*** (0.002)	0.007*** (0.001)	0.007*** (0.001)
Lost R-3	0.012*** (0.001)	0.011*** (0.001)	0.008*** (0.002)	0.008*** (0.002)	0.004*** (0.001)	0.003** (0.001)
Lost R-1 x Female		-0.049*** (0.006)		-0.050*** (0.011)		-0.012*** (0.004)
Lost R-2 x Female		-0.004 (0.003)		0.006 (0.005)		0.002 (0.003)
Lost R-3 x Female		0.002 (0.003)		0.001 (0.005)		0.006** (0.003)
Observations	278,887	278,887	278,887	278,887	278,887	278,887
R ²	0.129	0.130	0.133	0.133	0.078	0.078

Note: Dependent variable is a binary variable indicating whether a player interrupted their tournament play. All regressions include controls for the player’s own rating, the rating difference to their opponent, their age, their opponent’s title and the gender signaled in the opponent’s profile picture. All regressions use player fixed effects and tournament \times round fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.4 Waiting Time

B.4.1 Poisson Regression

As the dependent variable is the time elapsed until the next observed entry, a poisson-based model is suited as a regression model. The drawbacks of the model are that coefficients lend themselves less well to interpretation, as the marginal effects depend on the current value of the all independent variables. The sign of the coefficients however tells us the direction of the effect. The full model specification is as follows:

$$Days\ to\ Reenter_{it} = \exp(\alpha_i + \delta TourPerf_{it} + \gamma TourPerf_{it} \times Female_i + \theta X_{it} + tournament_t) + \epsilon_{it} \quad (B.3)$$

A comparison between the model in the main paper and the model presented here show that the direction of the performance measures (FIDE performance, percentage scores, and quartiles) are the same. Strong performers are quicker to reenter into competition than those with a weaker performance, both in terms of subjective and relative measures. There is weak evidence that a strong performance reduces the waiting time between tournaments more strongly for women than for men.

Table B.12: Poisson Regression - Days Until Reentry

	(1)	(2)	(3)	(4)	(5)	(6)
Percentage Score	-0.951*** (0.058)	-0.876*** (0.061)				
Percentage Score \times Female		-0.717*** (0.186)				
FIDE Perf. – Elo (in 100s)			-0.054*** (0.004)	-0.052*** (0.004)		
FIDE Perf. – Elo (in 100s) \times Female				-0.021* (0.012)		
Placement Quartile						
1 st Quartile					-0.459*** (0.034)	-0.443*** (0.036)
2 nd Quartile					-0.372*** (0.030)	-0.362*** (0.033)
3 rd Quartile					-0.254*** (0.027)	-0.240*** (0.030)
Placement Quartile \times Female						
1 st Quartile \times Female						-0.179* (0.094)
2 nd Quartile \times Female						-0.074 (0.077)
3 rd Quartile \times Female						-0.093 (0.072)
Observations	52,019	52,019	52,019	52,019	52,019	52,019
Adjusted R ²	0.389	0.390	0.387	0.387	0.386	0.386

Note: Dependent variable is how many days elapsed to the next tournament entry. All regressions include additional controls for the player’s rating at the day of the reentry choice and indicators whether they won a cash prize in their last tournament appearance. All regressions include player and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.4.2 Reentry Next Week

Instead of focusing on the number of days until a player reemerges in the tournament, this analysis checks whether a player’s performance is linked to their probability of appearing in the next possible tournament. The dependent variable is binary, indicating whether a player that played in tournament t will enter into tournament $t + 1$. The regression model follows Equation 12 with the same performance metrics as in the main paper and is formalized below:

$$Entered_{it+1} = \alpha_i + \delta TourPerf_{it} + \gamma TourPerf_{it} \times Female_i + \theta X_{it} + tournament_t + \epsilon_{it} \quad (B.4)$$

The results reported in Table B.13 showcase the same pattern that strong performances are associated with a higher likelihood of entering a tournament the following week. In particular the quartile analysis showcases how the estimated probability is reduced with lower placements. The results indicate that women may even respond more strongly to a good tournament performance than men do.

Table B.13: Likelihood of Next Possible Tournament Entry

	(1)	(2)	(3)	(4)	(5)	(6)
Percentage Score	0.242*** (0.014)	0.228*** (0.015)				
Percentage Score \times Female		0.136*** (0.041)				
FIDE Perf. – Elo (in 100s)			0.014*** (0.001)	0.014*** (0.001)		
FIDE Perf. – Elo (in 100s) \times Female				0.007** (0.003)		
Placement Quartile						
1 st Quartile					0.124*** (0.008)	0.122*** (0.008)
2 nd Quartile					0.101*** (0.007)	0.100*** (0.008)
3 rd Quartile					0.058*** (0.006)	0.055*** (0.007)
Placement Quartile \times Female						
1 st Quartile \times Female						0.008 (0.025)
2 nd Quartile \times Female						0.008 (0.020)
3 rd Quartile \times Female						0.016 (0.017)
Observations	55,364	55,364	55,364	55,364	55,364	55,364
Adjusted R ²	0.220	0.220	0.218	0.219	0.219	0.219

Note: Dependent variable is whether a player entered next week’s Titled Tuesday tournament after their appearance. All regressions include additional controls for the player’s rating at the day of the reentry choice and indicators whether they won a cash prize in their last tournament appearance. All regressions include player and tournament fixed effects. Standard errors are clustered at the individual level and are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C Online Appendix

C.1 Tournament Overview

Table C.1: List of Tournaments in Dataset

Date	Event	Games	No. Players
April 7, 2020	<i>Titled Tuesday Blitz</i>	1,889	455
April 14, 2020	<i>Titled Tuesday Blitz</i>	2,086	472
April 21, 2020	<i>Titled Tuesday Blitz</i>	2,010	465
April 28, 2020	<i>Titled Tuesday Blitz</i>	2,057	477
May 5, 2020	<i>Titled Tuesday Blitz</i>	1,739	404
May 12, 2020	<i>Titled Tuesday Blitz</i>	1,799	423
May 19, 2020	<i>Titled Tuesday Blitz</i>	1,029	242
May 26, 2020	<i>Titled Tuesday Blitz</i>	1,404	338
June 2, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,439	336
June 9, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,529	351
June 16, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,561	370
June 23, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,751	387
June 30, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,359	322
July 7, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,329	333
July 14, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,252	303
July 21, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,281	320
August 4, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,169	280
August 11, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,056	253
August 25, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,214	287
September 1, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,214	293
September 8, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,048	253
September 15, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,163	286
September 22, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,034	254
September 29, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	1,193	284
October 6, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	3,289	872
October 13, 2020	<i>SCC Grand Prix: Titled Tuesday Blitz</i>	4,190	1103
October 20, 2020	<i>Titled Tuesday Blitz</i>	3,736	1007
October 27, 2020	<i>Titled Tuesday Blitz</i>	4,060	1087
November 3, 2020	<i>Titled Tuesday Blitz</i>	4,203	1036
November 10, 2020	<i>Titled Tuesday Blitz</i>	4,254	1063
November 17, 2020	<i>Titled Tuesday Blitz</i>	3,955	986
November 24, 2020	<i>Titled Tuesday Blitz</i>	3,602	898
December 1, 2020	<i>Titled Tuesday Blitz</i>	4,244	1031
December 15, 2020	<i>Titled Tuesday Blitz</i>	2,879	718
December 22, 2020	<i>Titled Tuesday Blitz</i>	3,574	865
December 29, 2020	<i>Titled Tuesday Blitz</i>	3,045	752
January 12, 2021	<i>Titled Tuesday Blitz</i>	3,321	818
January 19, 2021	<i>Titled Tuesday Blitz</i>	3,616	869
January 26, 2021	<i>Titled Tuesday Blitz</i>	3,478	862
February 2, 2021	<i>Titled Tuesday Blitz</i>	3,004	735
February 9, 2021	<i>Titled Tuesday Blitz</i>	3,573	862

February 16, 2021	<i>Titled Tuesday Blitz</i>	3,064	745
February 23, 2021	<i>Titled Tuesday Blitz</i>	2,813	713
March 2, 2021	<i>Titled Tuesday Blitz</i>	3,069	765
March 9, 2021	<i>Titled Tuesday Blitz</i>	2,817	707
March 16, 2021	<i>Titled Tuesday Blitz</i>	2,799	701
March 23, 2021	<i>Titled Tuesday Blitz</i>	2,815	701
March 30, 2021	<i>Titled Tuesday Blitz</i>	2,638	668
April 6, 2021	<i>Titled Tuesday Blitz</i>	2,569	630
April 13, 2021	<i>Titled Tuesday Blitz</i>	2,531	627
April 20, 2021	<i>Titled Tuesday Blitz</i>	3,019	685
April 27, 2021	<i>Titled Tuesday Blitz</i>	2,545	597
May 4, 2021	<i>Titled Tuesday Blitz</i>	3,009	694
May 11, 2021	<i>Titled Tuesday Blitz</i>	3,055	697
May 18, 2021	<i>Titled Tuesday Blitz</i>	3,286	758
May 25, 2021	<i>Titled Tuesday Blitz</i>	2,952	689
June 1, 2021	<i>Titled Tuesday Blitz</i>	3,119	718
June 8, 2021	<i>Titled Tuesday Blitz</i>	2,986	672
June 15, 2021	<i>Titled Tuesday Blitz</i>	3,139	752
June 22, 2021	<i>Titled Tuesday Blitz</i>	3,062	707
June 29, 2021	<i>Titled Tuesday Blitz</i>	3,311	760
July 6, 2021	<i>Titled Tuesday Blitz</i>	3,216	732
July 13, 2021	<i>Titled Tuesday Blitz</i>	3,207	737
July 20, 2021	<i>Titled Tuesday Blitz</i>	3,489	793
July 27, 2021	<i>Titled Tuesday Blitz</i>	3,103	723
August 3, 2021	<i>Titled Tuesday Blitz</i>	3,154	709
August 10, 2021	<i>Titled Tuesday Blitz</i>	3,183	726
August 17, 2021	<i>Titled Tuesday Blitz</i>	3,356	767
August 24, 2021	<i>Titled Tuesday Blitz</i>	3,359	777
August 31, 2021	<i>Titled Tuesday Blitz</i>	1,644	391
September 7, 2021	<i>Titled Tuesday Blitz</i>	2,726	639
September 14, 2021	<i>Titled Tuesday Blitz</i>	3,108	725
September 28, 2021	<i>Titled Tuesday Blitz</i>	3,387	780
October 5, 2021	<i>Titled Tuesday Blitz</i>	3,217	749
October 12, 2021	<i>Titled Tuesday Blitz</i>	2,800	635
October 19, 2021	<i>Titled Tuesday Blitz</i>	2,589	597
October 26, 2021	<i>Titled Tuesday Blitz</i>	2,788	650
November 2, 2021	<i>Titled Tuesday Blitz</i>	2,720	641
November 9, 2021	<i>Titled Tuesday Blitz</i>	2,952	688
November 16, 2021	<i>Titled Tuesday Blitz</i>	2,517	582
November 23, 2021	<i>Titled Tuesday Blitz</i>	2,667	627
November 30, 2021	<i>Titled Tuesday Blitz</i>	2,629	611
December 7, 2021	<i>Titled Tuesday Blitz</i>	2,402	575
December 14, 2021	<i>Titled Tuesday Blitz</i>	2,509	576
December 21, 2021	<i>Titled Tuesday Blitz</i>	2,053	470
December 28, 2021	<i>Titled Tuesday Blitz</i>	2,190	506
January 4, 2022	<i>Titled Tuesday Blitz</i>	2,112	493
January 11, 2022	<i>Titled Tuesday Blitz</i>	2,033	461
January 18, 2022	<i>Titled Tuesday Blitz</i>	2,083	474
January 25, 2022	<i>Titled Tuesday Blitz</i>	2,022	465

February 1, 2022	<i>Titled Tuesday Blitz</i>	2,273	524
February 1, 2022	<i>Titled Tuesday Blitz</i>	2,358	563
February 8, 2022	<i>Early Titled Tuesday Blitz</i>	2,307	524
February 8, 2022	<i>Late Titled Tuesday Blitz</i>	2,338	543
February 15, 2022	<i>Early Titled Tuesday Blitz</i>	2,040	482
February 15, 2022	<i>Late Titled Tuesday Blitz</i>	1,990	474
February 22, 2022	<i>Early Titled Tuesday Blitz</i>	2,096	501
February 22, 2022	<i>Late Titled Tuesday Blitz</i>	2,045	490
March 1, 2022	<i>Early Titled Tuesday Blitz</i>	1,921	457
March 1, 2022	<i>Late Titled Tuesday Blitz</i>	1,889	447
March 8, 2022	<i>Early Titled Tuesday Blitz</i>	2,065	483
March 8, 2022	<i>Late Titled Tuesday Blitz</i>	1,839	436
March 15, 2022	<i>Early Titled Tuesday Blitz</i>	2,279	527
March 15, 2022	<i>Late Titled Tuesday Blitz</i>	2,121	506
March 22, 2022	<i>Early Titled Tuesday Blitz</i>	2,327	546
March 22, 2022	<i>Late Titled Tuesday Blitz</i>	2,152	504
March 29, 2022	<i>Early Titled Tuesday Blitz</i>	1,956	468
March 29, 2022	<i>Late Titled Tuesday Blitz</i>	2,447	579
April 5, 2022	<i>Early Titled Tuesday Blitz</i>	1,826	469
April 5, 2022	<i>Late Titled Tuesday Blitz</i>	1,573	361
April 12, 2022	<i>Early Titled Tuesday Blitz</i>	2,728	641
April 12, 2022	<i>Late Titled Tuesday Blitz</i>	2,485	612
April 19, 2022	<i>Early Titled Tuesday Blitz</i>	2,400	575
April 19, 2022	<i>Late Titled Tuesday Blitz</i>	2,514	586

Note: The table presents an overview of the individual tournaments that were used to construct the dataset.

C.2 Characteristics of Identified Individuals

Table C.2: Characteristics of Identified Accounts and Players

	Mean	SD	Min	Max	N
Female	0.1462	0.3533	0	1	3,523
Birthyear	1,989.3991	12.2993	1,940	2,011	3,523
Player Title:					
Title: GM	0.1805	0.3847	0	1	3,523
Title: IM	0.2047	0.4035	0	1	3,523
Title: FM	0.2884	0.4531	0	1	3,523
Title: CM	0.0917	0.2886	0	1	3,523
Title: NM	0.1110	0.3142	0	1	3,523
Title: WGM	0.0230	0.1499	0	1	3,523
Title: WIM	0.0349	0.1836	0	1	3,523
Title: WFM	0.0483	0.2143	0	1	3,523
Title: WCM	0.0173	0.1305	0	1	3,523
Title: WNM	0.0003	0.0168	0	1	3,523
Profile Picture:					
Profile Pic. - Male	0.5512	0.4974	0	1	3,523
Profile Pic. - None	0.2018	0.4014	0	1	3,523
Profile Pic. - Other	0.1493	0.3564	0	1	3,523
Profile Pic. - Female	0.0976	0.2969	0	1	3,523

Note: The table presents an overview of the characteristics of the identified players whose accounts could be linked to a FIDE profile.

The identification process exploits the clear names policy that chess.com has for the Titled Tuesday tournament. In order to be eligible for prizes, players must have their names listed in their user profiles if it is not evident from the username itself. A majority of players do this, alternatively they provide links to their national or FIDE player card or other social media which allows them to be reliably linked to their over-the-board record. Positive matches are further verified: first the provided name needs to match, additionally the title listed on the chess.com profile must be a title that the player won at some point in their career. If there is additional uncertainty, the flag that players can select on their chess.com profiles is matched as well.

Because of accounts closing, possible changes of account names, or choosing to no longer display the title the API was not able to find an entry for 68 accounts/players. As carrying a title is a requirement to participate in the tournament and this verification occurs manually, I impute the verified titles of 9 players I was able to retroactively identify. An additional 17 of 68 missing players could be identified based on their usernames but their titles could not be verified. The remaining players are mostly Brazilian and likely national masters, a title conferred by national chess federations instead of the international chess federation. The national master title cannot be separately verified on Brazilian Chess Federation's website, so these observations are dropped.

C.3 Extended Descriptive Statistics of Observed Games

Table C.3: Full Game-Level Descriptive Statistics

	Mean	SD	Min	Max	Observations
Average CPL	71.953	61.525	0	779	312,892
Aggressive Opening	0.267	0.442	0	1	203,646
Quit After Round	0.058	0.233	0	1	278,957
RoundScore	0.507	0.475	0	1	312,892
Number of Moves	44.632	17.465	10	301	312,892
Female	0.144	0.351	0	1	312,892
Age	31.709	12.775	10	82	312,892
Age ²	1,168.651	956.603	100	6,724	312,892
Player Elo (in 100s)	25.064	2.467	1	33	312,892
Elo Difference (in 100s)	0.065	2.171	-26	26	312,892
Lost R-1	0.413	0.492	0	1	312,892
Lost R-2	0.408	0.491	0	1	312,892
Lost R-3	0.402	0.490	0	1	312,892
Lost R-1 x Female	0.073	0.261	0	1	312,892
Lost R-2 x Female	0.075	0.263	0	1	312,892
Lost R-3 x Female	0.076	0.265	0	1	312,892
Upset Loss R-1	0.112	0.315	0	1	312,892
Upset Loss R-2	0.104	0.305	0	1	312,892
Upset Loss R-3	0.094	0.292	0	1	312,892
Upset Loss R-1 x Female	0.012	0.109	0	1	312,892
Upset Loss R-2 x Female	0.011	0.103	0	1	312,892
Upset Loss R-3 x Female	0.009	0.095	0	1	312,892
Expected Loss R-1	0.301	0.459	0	1	312,892
Expected Loss R-2	0.304	0.460	0	1	312,892
Expected Loss R-3	0.308	0.462	0	1	312,892
Expected Loss R-1 x Female	0.061	0.240	0	1	312,892
Expected Loss R-2 x Female	0.064	0.245	0	1	312,892
Expected Loss R-3 x Female	0.067	0.250	0	1	312,892
Opponent Title: GM	0.208	0.406	0	1	312,892
Opponent Title: IM	0.203	0.402	0	1	312,892
Opponent Title: FM	0.283	0.451	0	1	312,892
Opponent Title: CM	0.093	0.291	0	1	312,892
Opponent Title: NM	0.109	0.311	0	1	312,892
Opponent Title: WGM	0.020	0.139	0	1	312,892
Opponent Title: WIM	0.030	0.172	0	1	312,892
Opponent Title: WFM	0.040	0.195	0	1	312,892
Opponent Title: WCM	0.014	0.119	0	1	312,892
Opponent Title: WNM	0.000	0.016	0	1	312,892
Opp Profile Pic. - Male	0.521	0.500	0	1	312,892
Opp Profile Pic. - None	0.225	0.418	0	1	312,892
Opp Profile Pic. - Other	0.168	0.374	0	1	312,892
Opp Profile Pic. - Female	0.085	0.279	0	1	312,892

Note: The table presents an overview of the game-level descriptive data for the analytical sample.

C.4 Extended Descriptive Statistics of Waiting Time

Table C.4: Full Tournament-Level Descriptive Statistics

	Mean	SD	Min	Max	N
Weeks to Tournament Reentry	3.329	7.259	0	102	49,927
Entered Next Week	0.582	0.493	0	1	49,927
Percentage Score	0.418	0.212	0	1	49,927
7 Day Sum of Blitz Games	30.345	50.030	0	775	49,927
7 Day Sum of Bullet Games	11.653	50.231	0	1,390	49,927
7 Day Sum of Rapid Games	0.638	4.373	0	315	49,927
FIDE Performance Measure	2,397.478	316.958	1,016	3,617	49,927
Tournament Elo Average (in 100s)	25.071	2.448	11	33	49,927
FIDE Perf. - Elo	-109.615	254.270	-1,625	1,209	49,927
1 st Quartile	0.285	0.451	0	1	49,927
2 nd Quartile	0.246	0.431	0	1	49,927
3 rd Quartile	0.241	0.428	0	1	49,927
4 th Quartile	0.228	0.420	0	1	49,927
1 st Quartile \times Female	0.016	0.127	0	1	49,927
2 nd Quartile \times Female	0.033	0.180	0	1	49,927
3 rd Quartile \times Female	0.042	0.201	0	1	49,927
4 th Quartile \times Female	0.039	0.194	0	1	49,927
Elo Change from First to Last Round	0.500	29.204	-419	2,033	49,927
Sonneborn-Berger Tiebreak Score	19.443	13.830	0	83	49,927
Share of Rounds Played	0.810	0.276	0	1	49,927
Age in Years	31.242	12.606	10	82	49,927
Age ²	1,134.980	939.021	100	6,724	49,927
1st place	0.002	0.046	0	1	49,927
2nd place	0.002	0.046	0	1	49,927
3rd place	0.002	0.047	0	1	49,927
4th place	0.002	0.047	0	1	49,927
5th place	0.002	0.046	0	1	49,927
Best Woman	0.002	0.045	0	1	49,927

Note: The table presents an overview of the tournament-level descriptive data for the analytical sample.