I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13491

Cognitive Performance in the Home Office - Evidence from Professional Chess

Steffen Künn
Christian Seel
Dainis Zegners

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13491

Cognitive Performance in the Home Office - Evidence from Professional Chess

Steffen Künn
Maastricht University, IZA and ROA

Christian Seel
Maastricht University

Dainis Zegners
Erasmus University

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Cognitive Performance in the Home Office - Evidence from Professional Chess

During the recent COVID-19 pandemic, traditional (offline) chess tournaments were prohibited and instead held online. We exploit this as a unique setting to assess the impact of moving offline tasks online on the cognitive performance of individuals. We use the Artificial Intelligence embodied in a powerful chess engine to assess the quality of chess moves and associated errors. Using within-player comparisons, we find a statistically and economically significant decrease in performance when competing online compared to competing offline. Our results suggest that teleworking might have adverse effects on workers performing cognitive tasks.

JEL Classification: H12, L23, M11, M54

Keywords: teleworking, productivity, chess, COVID-19

Corresponding author:
Christian Seel
Maastricht University
Department of Microeconomics and Public Economics
Tongersestraat 53
6211 LM
The Netherlands
E-mail: c.seel@maastrichtuniversity.nl

1 Introduction

Teleworking (also known as telecommuting or working from home) has seen a steep increase during the recent COVID-19 pandemic. In a recent survey, half of U.S. workers reported working from home during the pandemic in April and May 2020 (Brynjolfsson et al., 2020). While the jump has been driven both by voluntary and mandated social distancing, it is arguably an acceleration of a broader trend towards more flexible work arrangements (Mas and Pallais, Forthcoming) and more outsourcing enabled by digital technologies (Agrawal et al., 2015) increasing the number of workers working from home. Dingel and Neiman (2020) estimate that 37% of jobs in the U.S. could be done entirely from home.

An important question for firms and regulators is how this trend towards more teleworking affects workers' productivity. Yet, despite the large societal relevance, the literature in economics on the topic is sparse. A major hurdle for empirical work is to isolate changes in the type of work and tasks that workers perform when working from home from changes in individual productivity. We contribute towards filling this gap by analyzing the performance of professional chess players who compete in chess tournaments that are organized online and offline but that are otherwise conducted under comparable conditions.

The analysis is based on comparing the performance of elite professional chess players competing in a recently organized *online* tournament to their performance during recent *offline* tournaments. During the COVID-19 pandemic when physical contact among players was prohibited, the current world champion Magnus Carlsen initiated an online event, the *Magnus Carlsen Invitational*. We use this event to compare the performance of the participating players to their performance in recent editions of the *World Rapid Chess Championship* as organized by the World Chess Federation in a traditional offline setting. Both tournaments are organized under comparable conditions, in particular giving players the same amount of thinking time during a game, and offer comparable prize funds. Our benchmark of performance is based on evaluating the moves played by the participants using a currently leading chess engine, which significantly outperforms the best human players in terms of playing strength.

Comparing online with offline chess tournaments offers several advantages for assessing the impact of moving tasks online on cognitive performance. First, playing chess is a purely cognitive task, which requires complex strategic decision making under time constraints. Therefore, chess offers a unique setting for studying performance in a cognitive task, which is important in many modern professional, managerial, technical, and creative occupations (Autor and Price, 2013). Second, although until very recently high-stakes chess tournaments were almost exclusively conducted with players competing face-to-face in physical playing halls,[1] most chess players are very familiar with unincentivized online chess on various chess platforms. Due to the recent COVID-19 pandemic, several online tournaments are being organized in which many of the world's elite players are participating, usually playing from their homes. These tournaments offer significant amounts of prize money to the players providing them with high incentives for performance. Third, using the Artificial Intelligence embodied in modern chess engines makes it possible to construct a benchmark of individual performance that is based on fine-grained move-by-move data with a high degree of objectivity and accuracy. This benchmark makes it possible to analyze both the probability and magnitude of making mistakes during a chess game. Fourth, since all major global chess events were canceled during the pandemic, we are able to observe a representative sample of elite players competing both online and offline, ruling out selection-effects influencing our results.

Analyzing 27,267 individual moves played during 441 games in a regression model with player fixed effects, we provide evidence for a statistically and economically significant

[1] The lack of official online tournaments with significant prize money is mainly due to the potential for cheating by using a chess engine as even a chess engine running on a mobile device vastly surpasses the human World Champion in terms of playing strength. Even in larger online tournaments with only a couple of hundred dollars in prizes, there are frequent allegations of cheating, which is difficult to detect.

decrease in performance when the same players compete online compared to competing offline. While the probability of playing the best move as suggested by the chess engine is not statistically different online and offline, we find that conditionally on making an error, the magnitude of the error is 16.8% larger online for the sample player. This difference in performance measured in terms of error size is statistically significant at the 1% level.

We contribute to the literature examining the impact of teleworking (also called telecommuting) and working from home on workers' productivity. A large body of studies in psychology uses mostly unincentivized survey data. In a meta-study, Gajendran and Harrison (2007) find no effect on self-reported performance and a positive effect of teleworking on supervisor-reported or archival records of performance. As a conclusion, they state that "A common refrain in reviews of telecommuting research has been the inability, over 20 years of studies, to draw consistent conclusions about even its most basic consequences" (p. 1538). A lack of clear evidence on the effect of telecommuting on productivity is also reported in other literature surveys, e.g., Bailey and Kurland (2002) and Allen et al. (2015).

The seminal paper in the economics literature is Bloom et al. (2015), who examine the productivity of call-center workers in a randomized controlled trial. They find positive effects of working from home on productivity that are driven by higher effort (more minutes per shift and fewer sick days) and effectiveness (more calls per minute due to a better work environment). They also examine conversion rates and externally evaluated call quality, and did not find statistically significant effects.

Our study complements the study of Bloom et al. (2015). In contrast to them, we consider a highly specialized cognitively demanding task. In our setting, we can directly measure performance using an Artificial Intelligence based measure instead of a proxy such as effort or effectiveness. This allows us to estimate changes in individual productivity that are due to working from home and are purely driven by task-level cognitive performance.

There are a few other related recent studies. Using public sector data, Linos (2016) finds in a within-subject design that teleworking patent officers have a lower productivity per hour, but make up for it by spending a larger portion of their workday on their core task and less time in meetings. Angelici and Profeta (2020) find increases in objective worker productivity in a knowledge firm in which workers are randomized into a treatment that allows for more flexible work arrangements in terms of hours worked and location. In a lab experiment, Dutcher (2012) simulates a dull work task (typing numbers and letters on a computer keyboard, mimicking data entry) and a creative task (playing tic-tac-toe against a computer). He finds a positive impact on the creative task of conducting it online and a negative impact for the dull task.

The literature discusses several potential reasons for productivity differences at home compared to the office environment. Many employers fear shirking from home due to distractions while Beckmann (2016) (p.8) claims that for the call-center employees in the Bloom et al. (2015) study, there is a "scope for productivity enhancements because employees working in large and noisy offices were easily distracted." In our study, this channel does not play a major role, as noise levels are low online and offline. In addition, players were constantly monitored by webcams and highly incentivized to focus on their task.

A crucial difference to the offline setting is that the peer pressure to concentrate in a playing hall is missing. For instance, Falk and Ichino (2006) find that students place letters in envelopes at a higher speed when other students are faced with the same task sit in the room. Finally, in general, players might have a dip in performance as either teleworking or the pandemic could cause a reduction in the general mental well-being; e.g., Bloom et al. (2015) report an increased feeling of loneliness among teleworkers.

2 Data and Methods

We use chess as an empirical setting to study cognitive performance. Playing chess is a complex, strategic, and cognitively demanding task that has been heavily used by cognitive psychologists to investigate strategic and cognitive aspects of human thinking, such as perception, memory, and problem-solving (e.g. de Groot, 1946; Chase and Simon, 1973; Simon and Chase, 1973; Charness, 1992). Burgoyne et al. (2016) survey the empirical evidence for the relationship between chess skill and general cognitive skills such as fluid reasoning, comprehension knowledge, short-term memory, and processing speed. In recent years, economists have used chess to examine questions related to rationality (Palacios-Huerta and Volij, 2009; Levitt et al., 2011; González-Díaz and Palacios-Huerta, 2016; Zegners et al., 2020), gender (Gerdes and Gränsmark, 2010; Backus et al., 2016), adverse effects of pollution (Künn et al., 2019), and age (Bertoni et al., 2015; Strittmatter et al., 2020).

2.1 Data Collection

Our data consist of games from the *World Rapid Chess Championships* 2015 - 2019 played offline in Berlin, Doha, Riyadh, Saint Petersburg and Moscow and the *Magnus Carlsen Invitational* tournament played online from April 18, 2020 till May 3, 2020 on the Internet chess platform *chess24.com*. The selected tournaments are identical with respect to the time limit as players are given a time budget of 15 minutes to complete the game with 10 seconds added to a player's time budget for each move played. In contrast to shorter Blitz games (usually 3-5 minutes time limit per player), small differences in the time of physically executing a move and pressing the clock as compared to entering the move to a computer are unlikely to have a strong impact on the outcome in relatively longer rapid games. Finally, the majority of players in the *Magnus Carlsen Invitational* also competed in at least one edition of the *World Rapid Chess Championships* enabling us to make within-player comparisons of performance.

The *World Rapid Chess Championships* offered an overall prize pool of $200,000 in 2015 and 2016, $750,000 in 2017 and $350,000 in 2018 and 2019 to the participating players. These included more than a hundred players among them many of the world's elite players. The tournament format was a 15 round Swiss tournament, i.e., players with similar rankings in the tournament standing are paired against each other in each round, but the same opponents can only play each other once. The winner was the player with the highest score out of 15 games.[2] To prevent cheating, there were certified walk-through metal detectors at the entrance of the playing hall.

The *Magnus Carlsen Invitational* offered a prize pool of $250,000 to the participating players. These included eight players who are among the world's elite and are ranked between 1 and 21 in the official "FIDE World ranking" for classical chess. The tournament differs from other online tournaments in terms of the strict anti-cheating measures that included arbiters monitoring players at all times and standard automated cheating detection systems in place.[3] Moreover, several commentators agreed that given their high standing in the world rankings, players would be very careful to avoid any suspicion of cheating as this would greatly damage their reputation.[4] The tournament was split into two phases, first a league and then a knockout phase. In the league, each player played a mini-match against each other participant. Each mini-match featured four games and the player who scored more points obtained 3 points in the league, while the loser received 0 points.[5] The top four players then qualified for the semi-finals. The format of the semi-finals was similar

[2]In the case of a tie, playoff of "Blitz" games in which players have a substantially smaller time budget took place to determine the World Rapid Chess Champion. We disregard such games from our analysis.

[3]These systems compare the moves played by a player to the optimal moves suggested by the leading chess engines, flagging a player as suspect of cheating if there is a too large agreement to the engine or thinking time patterns that are indicative of cheating.

[4]See for example `https://en.chessbase.com/post/magnus-carlsen-invitational-2020-preview` (accessed on June 10, 2020)

[5]In the case of a tie, an "Armageddon Blitz" game with a substantially smaller time budget was played to determine the winner of a match. We disregard these games from our analysis.

to a mini-match in the league, again with 4 games played with the same time budget for each player.[6] The winners of the semi-finals advanced to the finals, which were played under the same format as the semi-finals.

We include in our analysis all games that were played in the online tournament and all games from the offline tournaments in which one of the eight players from the online tournament participated. We further remove the opening phase for each game, defined as the first 15 moves for each player (as in Backus et al., 2016), because players usually play moves that they memorize as part of their preparation and training. In total, we observe 8,260 (19,007) moves played in 123 (318) games from the online (offline) tournament.

2.2 Evaluation of Chess Moves

To estimate the effect of playing online on chess players' performance, we evaluate each move in each chess game in our sample using a chess engine. We use the chess engine Stockfish 11 for this purpose, which during the last decade has been consistently ranked first or near the top among chess engines. Modern chess engines such as Stockfish 11 considerably outperform every human player on off-the-shelve computer hardware in terms of ELO rating, i.e., the method used by the World Chess Federation to measure the strength of a player.[7]

We assess the performance of players based on the amount and size of errors they make according to the evaluation of the chess engine. A chess game g consists of *moves* $m_g \in \{1, \ldots, M_g\}$, where a move consists of one *individual move* m_{ig} by each player i

[6] In the case of a tie, a playoff of games with a substantially smaller time budget was played to determine the winner of a match. As before, we disregard these games from our analysis.

[7] As of May 2020, Stockfish 11 is rated with an estimated ELO rating of 3494 and hence, clearly outperforms any human player. The ELO rating is a measure of relative chess strength introduced by the Hungarian mathematician Arpad Elo (Elo, 1978). As a comparison, the best current human player is Magnus Carlsen (also included in our sample) who has an ELO rating of 2863. The player with the lowest rating in the online tournament in our sample has an ELO rating of 2728. See the unofficial rating list for chess engines at `http://www.computerchess.org.uk/ccrl` and the official ELO rating list published by the World Chess Federation at `https://ratings.fide.com` (both accessed on May 27, 2020).

(the last move M_g might only feature one individual move by the player who has the White pieces). For a given position of game g before individual move m_{ig} the chess engine computes an evaluation of the position in terms of the *pawn metric* P_{igm}. As chess is a zero-sum game, the advantage of one player is equal to the disadvantage of the other player, where $P_{igm} > 0$ ($P_{igm} < 0$) indicates an advantage (disadvantage) for player i. The numerical value of the pawn metric indicates the size of the advantage from the perspective of player i, with one unit indicating an advantage that is comparable to being one pawn up.[8] The pawn metric is computed assuming that both players play optimal moves, i.e., the game proceeds along the optimal path computed by the chess engine.[9] For each player i in each game g at each move m_{ig}, we compute two pawn metrics: $\overline{P}_{igm}$ denoting the pawn metric before player i makes his move and $\underline{P}_{igm}$ denoting the pawn metric of the chess engine after player i makes his move. Using these two measures, we compute for each move an error defined as

$$Error_{igm} = \overline{P}_{igm} - \underline{P}_{igm}, \tag{1}$$

which reflects the change in the pawn metric after player i has made his move m_{ig}.

Intuitively, the $Error_{igm}$ variable should be zero after an optimal move and positive after a non-optimal move. Yet, there is a small amount of randomness in the evaluation function, which we will account for with a random error term in our regression and in a separate robustness analysis.[10] We provide an example of the output of the chess engine

[8]Other characteristics of a chess position that are relevant for assessing a player's winning chances such as having a weak King's position or a good pawn-structure are also factored-in into the pawn-metric. See `https://chess.fandom.com/wiki/Centipawn` (accessed on June 16, 2020).

[9]The chess engine starts with the current position as the root of a game tree. It then builds the game tree for a pre-specified number of moves in the tree ahead (the search depth) using an alpha-beta pruning algorithm with iterative deepening (based on good play by both sides) and a transposition table. It assigns positions at the terminal nodes of the tree a value using an evaluation function. For more details, see `http://rin.io/chess-engine/`, last visited June 16, 2020). We restrict STOCKFISH 11 to a search depth of 25 moves ahead to economize on computing costs.

[10]There are two sources of randomness: 1) As the engine is set to calculate the game tree arising from

and the computation of the error metric in Figure A.1 in the appendix.

In addition to the evaluation of the position, the chess engine returns the number of unique nodes of the game tree that it had to search to reach a pre-specified search depth. We use this information as a measure of the complexity of the position, as it is directly related to the branching factor of the game tree for a given chess position.

2.3 Estimation Strategy and Outcome Variables

To estimate the impact of playing online on a player's performance, we estimate the following linear model:

$$Y_{igm} = \alpha + \delta Online_g + \beta X_{igm} + \eta_i + \gamma_m + V_{igm}, \tag{2}$$

where Y_{igm} is the outcome variable measured in game g played by player i at move m.

The term $Online_g$ denotes the treatment indicator taking the value one if game g was played in the online tournament *Magnus Carlsen Invitational* and zero otherwise. Our parameter of interest is denoted by δ, which measures the difference in outcome variables between games conducted online and offline. We identify the parameter of interest by observing the same individuals i playing moves in the online and the offline tournaments.[11]

Our regression model includes the following set of time-, game- and move-specific controls that are included in vector X_{igm}: (i) A measure representing the complexity of the position before the player makes his move, (ii) the current *ELO* rating of the player to

a position to a pre-specified search depth, it will calculate one move deeper in the position after a player has made a move. 2) To save on computing time, a chess engine does not evaluate branches of the game tree that it has found to be dominated by another branch. This creates an effect whereby the randomly determined search-order of moves has a small impact on the final evaluation of a position. The impact of these sources is empirically small in size, centered around zero, and unlikely to be correlated with other characteristics of the position and variables included in the regression model. Therefore, in our view, they will be sufficiently accounted for by the error term in our linear regression model.

[11] Our data also includes the moves of the opponents in the offline tournaments. Due to the individual player fixed effects, however, these players do not contribute towards estimating the main effect of playing online.

move as well as the difference in the ELO rating to the opponent,[12] (iii) the number of games played before game g within the tournament as well as during a specific day, (iv) the remaining time budget of a player before each move, and (v) the absolute value of the current pawn metric of the position before the player makes his move $\overline{P}_{igm}$ as well as its squared term. η_i and γ_m are individual player and move fixed effects, respectively. Finally, the error term V_{igm} is clustered at the game level to allow for arbitrary correlation within each game.

Although we cannot make final statements concerning causality because of the absence of an experimental setting, the rich specification makes us very confident that δ is likely to represent the causal parameter of playing online (vs. offline) on outcome variables.

We use the following outcome variables that are constructed based on the raw error measure in Eq. (1). The first is a binary transformation such that

$$Make\,Error_{igm} = \begin{cases} 1 & \text{if } Error_{igm} > 0 \\ 0 & \text{if } Error_{igm} \leq 0, \end{cases} \tag{3}$$

which indicates whether the move played decreases the pawn metric and thus is an error.

The second transformation uses the logarithm of the error of the form

$$Ln\,Error_{igm} = \begin{cases} ln(Error_{igm}) & \text{if } Make\,Error_{igm} = 1 \\ 0 & \text{otherwise}, \end{cases} \tag{4}$$

which measures the magnitude of errors conditional on an error being conducted.

[12] We use the official ELO rankings by the World Chess Federation for rapid chess, see `https://ratings.fide.com/top_lists.phtml`.

3 Results

Table 1 contains our main estimation results and shows the estimated coefficient $\hat{\delta}$ based on equation 2. Each row presents the results of a separate regression using the different outcome variables as explained in Section 2.3.

Table 1 Main results: Offline vs. online tournament setting on performance of chess players

Outcome Variable	Number ind. moves	(1)	(2)	(3)
Make Error	27,267	0.010	-0.014	0.021
		(0.396)	(0.409)	(0.138)
Ln Error if Make Error = 1	15,173	0.136***	-0.001	0.168***
		(0.009)	(0.989)	(0.002)
Controls		YES	NO	YES
Player FE		NO	YES	YES
Move FE		NO	YES	YES

Note: The table shows the estimated coefficient $\hat{\delta}$ based on equation 2. Each row presents the results of a separate regression using different outcome variables. Standard errors are clustered at the game level and p-values are reported in parenthesis. Section 2.3 describes the construction of the outcome variables. The set of control variables includes: (i) a measure representing the complexity of the position in which the move was made, (ii) the current ELO rating of the player as well as the difference in the ELO rating to the opponent, (iii) the number of games played before game g within the tournament as well as during a specific day, (iv) the remaining time before each move, and (v) the absolute value of the current pawn metric of the position before the player makes his move $\overline{P}_{igm}$ as well as its squared term. The opening phase of each game is excluded for each player ($m \leq 15$). Descriptive statistics of the included variables as well as full estimation results for the final specification (column 3) are shown in Table A.1 and Table A.2 in the appendix, respectively. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$

In the following, we discuss our preferred model using the full specification including all control variables and the full set of fixed effects as shown in column (3) in Table 1. Using making an error as an outcome variable, we find a positive coefficient on the treatment indicator (playing online) that is, however, not statistically significant at conventional levels (p-value=0.138). Conditional on making an error, we find that players make on average 16.8% larger errors when playing online. This effect is statistically significant at the 1%-

level indicating that the online setting induces a reduction in the performance of chess players that is driven by an increase in the magnitude of errors.

To better assess the size of the effect, we provide a back-of-the-envelope calculation for the change in playing strength when playing online as expressed in terms of the *ELO* rating. In our sample, the coefficient on the *ELO* rating of the player (-0.0007672, rounded to -0.001 in table A.2 in the appendix) indicates that if a player's *ELO* rating increases by one unit, the magnitude of the error is reduced by 0.077% on average. Playing online increases the error size by 16.8% which corresponds to a loss of 219 points of *ELO* rating. The factual drop in playing strength on a game level is likely to be lower. First, our analysis excludes the opening stage which is prepared and memorized by the players prior to the games. The quality of play in this part differs across players, but likely does not differ online and offline. Second, we use a linear regression for our calculation for the translation of error size to *ELO*. Yet, as error margins are smaller at the top, a further drop in error by the same percentage likely results in a higher gain in terms of *ELO* rating at the top.

We test the sensitivity of our results with respect to (i) alternating the definition of the opening phase, (ii) excluding moves in positions that are evaluated as $|\overline{P}_{igm}| > 2$ indicating that one player already faces a significant (dis)advantage potentially altering players' behavior, and (iii) applying a more restrictive definition of errors, i.e., only considering moves as errors with a change in the pawn metric larger than 0.1 and not being annotated by the chess engine as the best possible move. The latter should test whether our results are possibly driven by marginal or mechanical errors created by the randomness in the evaluation of the chess engine. Table 2 summarizes the results of the sensitivity analysis. First, when implementing a more restrictive definition of the opening phase (column 3) or applying a more restrictive definition of errors (column 5), the effect on the probability of making an error slightly increases in size compared to the main result (column 1) and becomes statistically significant at conventional levels. However, when excluding moves in positions that are evaluated as $|\overline{P}_{igm}| > 2$ (column 4), the effect on the probability to make an error disappears completely. This suggests that the effect on the probability of making an error seems to be driven by errors in positions that are already relatively (dis)advantageous for a player. In contrast, the effect on the size of the error is very robust with respect to all sensitivity checks.

Table 2 Sensitivity analysis

	Main results (see Table 1)	Excluding opening phase$^{a)}$ $m \leq 10$	$m \leq 20$	Excluding moves with $\lvert\overline{P}_{igm}\rvert > 2^{b)}$	Restrictive def. of errors$^{c)}$
	(1)	(2)	(3)	(4)	(5)
Make Error	0.021	0.020	0.026*	0.000	0.025**
	(0.138)	(0.117)	(0.092)	(0.983)	(0.034)
	[27,267]	[31,709]	[22,922]	[20,501]	[27,267]
Ln Error if Make Error = 1	0.168***	0.139***	0.188***	0.121**	0.145***
	(0.002)	(0.004)	(0.002)	(0.034)	(0.008)
	[15,173]	[18,092]	[12,308]	[11,028]	[7,805]
Controls	YES	YES	YES	YES	YES
Player FE	YES	YES	YES	YES	YES
Move FE	YES	YES	YES	YES	YES

Note: The table shows the estimated coefficient $\hat{\delta}$ based on equation 2. Each row presents the results of a separate regression using different outcome variables. Standard errors are clustered at the game level and p-values are reported in parenthesis. Number of individual moves are in brackets. Section 2.2 describes the construction of the outcome variables. The set of control variables includes: (i) a measure representing the complexity of the actual move, (ii) the current *ELO* rating score of the player as well as the difference in the *ELO* score to the opponent, (iii) the number of games played before game g within the tournament as well as during a specific day, (iv) the remaining time before each move, and (v) the absolute value of the current pawn metric of the position before the player makes his move $\overline{P}_{igm}$ as well as its squared term.

a) The estimation of the main results is based on a sample excluding the opening phase which is defined by the first 15 moves. Here, we show the sensitivity of the results by further restricting ($m \leq 10$) and extending ($m \leq 20$) this condition.

b) Excluding moves in positions with a pawn metric $|\overline{P}_{igm}| > 2$.

c) Regarding the error variable, we exclude erroneous moves with marginal errors between zero and 0.1, or moves being the best possible as indicated by the chess engine. *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$

Finally, we investigate potential effect heterogeneity with respect to (i) the strength of the players, (ii) the duration of the games, and (iii) the progress in the tournaments. By this, we test whether stronger players are more capable of playing online, and whether the negative effect is transitory and maybe mitigates over the duration of the game or the tournament. We include an interaction term between the *online* dummy and the variable of interest (*ELO* rating of the player, move number or the number of games played before game g within the tournament) in our main regression model as shown in Equation 2. We find no significant coefficients on the interaction terms indicating no effect heterogeneity within our estimation sample.[13]

4 Conclusion

In this paper, we have compared the performance of professional chess players when playing in traditional (offline) tournaments with their performance during a recent online tournament. The online tournament was organized during the COVID-19 pandemic when any physical contact between players was prohibited. This provides a unique setting to assess the potential impact of moving offline tasks online on the cognitive performance of individuals. Observing chess players has a number of advantageous features that allow us to identify the effect. First, players were executing the same (purely) cognitive task repeatedly under identical tournament rules. Second, we have an objective measure of individual performance by evaluating each move in our sample of games using a chess engine. Finally, all players in our sample faced strong incentives to exert high effort because of high monetary prices.

Applying a fixed effect strategy, we identify the effect of playing online on players' performance by comparing the quality of moves played during online and offline tournaments by the same player. Our results indicate a significant decrease in performance when playing online. In particular, while we do not find a statistically significant increase in the probability of making an error, the size of an error when playing online increases by 16.8%. Thus, the cognitive performance of chess players is impaired when playing online. This effect might be explained by missing peer pressure as well as the intense atmosphere

[13]Results are available upon request.

during offline chess tournaments. Unfortunately, we are not able to provide an in-depth consideration of underlying mechanisms due to the restricted dataset. Moreover, comparing online games played during the COVID-19 pandemic with offline games played before the pandemic, a possible concern is that the decrease in cognitive performance does not just capture the effect of teleworking but also other confounding factors related to the COVID-19 pandemic, such as uncertainty, anxiety, income loss etc. (Brodeur et al., 2020). While we cannot completely rule out this possibility, we believe that such factors play a negligible role in our setting given our focus on wealthy, highly trained, world-elite chess players. It is unlikely that these confounding factors affect the players' well-being and performance on the chessboard (see e.g. Papageorge et al., 2020, documenting a larger burden for individuals with lower incomes). Additionally, our sample includes players from a diverse set of countries, some of which experienced only mild outbreaks of the virus (e.g. Norway) or that had already successfully contained the outbreak (e.g. China).

During the recent COVID-19 lockdown, millions of workers had to adjust to a home office environment overnight, basically moving workers' tasks and communication completely online. Our results suggest that such an adjustment might have adverse effects on workers' performance on cognitive tasks. It remains to investigate whether this adverse effect on cognitive performance is rather transitory or permanent. People might adapt to online tasks in the long-run. We could not find evidence supporting the adaption hypothesis within the observed tournament which might be just because it captures a too short period (123 games played over a period of two weeks).

References

Agrawal, Ajay, John Horton, Nicola Lacetera, and Elizabeth Lyons, "Digitization and the contract labor market: A research agenda," in Avi Goldfarb, Shane Greenstein, and Catherine Tucker, eds., *Economic analysis of the digital economy*, University of Chicago Press, 2015, pp. 219–250.

Allen, Tammy D, Timothy D Golden, and Kristen M Shockley, "How effective is telecommuting? Assessing the status of our scientific findings," *Psychological Science in the Public Interest*, 2015, *16* (2), 40–68.

Angelici, Marta and Paola Profeta, "Smart-working: Work flexibility without constraints," 2020.

Autor, David and Brendan Price, "The changing task composition of the US labor market: An update of Autor, Levy, and Murnane (2003)," *MIT Working Paper*, 2013.

Backus, Peter, Maria Cubel, Matej Guid, Santiago Sanchez-Pages, and Enrique Mañas, "Gender, competition and performance: Evidence from real tournaments," *IEb working paper 2016/27*, 2016.

Bailey, Diane E and Nancy B Kurland, "A review of telework research: Findings, new directions, and lessons for the study of modern work," *Journal of Organizational Behavior*, 2002, *23* (4), 383–400.

Beckmann, Michael, "Working-time autonomy as a management practice," *IZA World of Labor*, 2016.

Bertoni, Marco, Giorgio Brunello, and Lorenzo Rocco, "Selection and the age-productivity profile. Evidence from chess players," *Journal of Economic Behavior and Organization*, 2015, *110*, 45–58.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying, "Does working from home work? Evidence from a Chinese experiment," *The Quarterly Journal of Economics*, 2015, *130* (1), 165–218.

Brodeur, Abel, Andrew E. Clark, Sarah Fl?che, and Nattavudh Powdthavee, "COVID-19, Lockdowns and Well-Being: Evidence from Google Trends," *IZA Discussion Paper*, 2020, *13204.*

Brynjolfsson, Erik, John J. Horton, Adam Ozimek, Garima Rock Daniel Sharma, and Hong-Yi TuYe, "COVID-19 and remote work: An early look at US Data," *NBER Working Paper No. 27344*, 2020.

Burgoyne, Alexander P, Giovanni Sala, Fernand Gobet, Brooke N Macnamara, Guillermo Campitelli, and David Z Hambrick, "The relationship between cognitive ability and chess skill: A comprehensive meta-analysis," *Intelligence*, 2016, *59*, 72–83.

Charness, Neil, "The impact of chess research on cognitive science," *Psychological Research*, 1992, *54* (1), 4–9.

Chase, William G and Herbert A Simon, "Perception in chess," *Cognitive psychology*, 1973, *4* (1), 55–81.

de Groot, Adrianus Dingeman, *Het denken van den schaker: Een experimenteel-psychologische studie*, Noord-Hollandsche Uitgevers Maatschappij Amsterdam, 1946.

Dingel, Jonathan I. and Brent Neiman, "How many jobs can be done at home?," *Journal of Public Economics*, 2020, *189.*

Dutcher, E Glenn, "The effects of telecommuting on productivity: An experimental examination. The role of dull and creative tasks," *Journal of Economic Behavior & Organization*, 2012, *84* (1), 355–363.

Elo, Arpad, "The rating of chess players, past and present (Arco, New York)," 1978.

Falk, Armin and Andrea Ichino, "Clean evidence on peer effects," *Journal of Labor Economics*, 2006, *24* (1), 39–57.

Gajendran, Ravi S and David A Harrison, "The good, the bad, and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences.," *Journal of Applied Psychology*, 2007, *92* (6), 1524.

Gerdes, Christer and Patrik Gränsmark, "Strategic behavior across gender: A comparison of female and male expert chess players," *Labour Economics*, 2010, *17* (5), 766–775.

González-Díaz, Julio and Ignacio Palacios-Huerta, "Cognitive performance in competitive environments: Evidence from a natural experiment," *Journal of Public Economics*, 2016, *139*, 40–52.

Künn, Steffen, Juan Palacios, and Nico Pestel, "Indoor air quality and cognitive performance," *IZA Discussion Paper*, 2019, *12632.*

Levitt, Steven D., John A. List, and Sally E. Sadoff, "Checkmate: Exploring backward induction among chess players," *American Economic Review*, 2011, *101* (2), 975–990.

Linos, Elizabeth, "Three essays on human capital in the public sector," *Doctoral dissertation, Harvard University, Graduate School of Arts and Sciences*, 2016.

Mas, Alexandre and Amanda Pallais, "Alternative work arrangements," *Annual Review of Economics*, Forthcoming.

Palacios-Huerta, Ignacio and Oscar Volij, "Field centipedes," *American Economic Review*, 2009, *99* (4), 1619–35.

Papageorge, Nicholas W., Matthew V. Zahn, Michele Belot, Eline van den Broek-Altenburg, Syngjoo Choi, Julian C. Jamison, and Egon Tripodi, "Socio-Demographic Factors Associated with Self-Protecting Behavior during the COVID-19 Pandemic," *IZA Discussion Paper*, 2020, *13333.*

Simon, Herbert A. and William G. Chase, "Skill in chess," *American Scientist*, 1973, *61* (4), 394–403.

Strittmatter, Anthony, Uwe Sunde, and Dainis Zegners, "Life cycle productivity patterns over the long run," *mimeo, LMU Munich*, 2020.

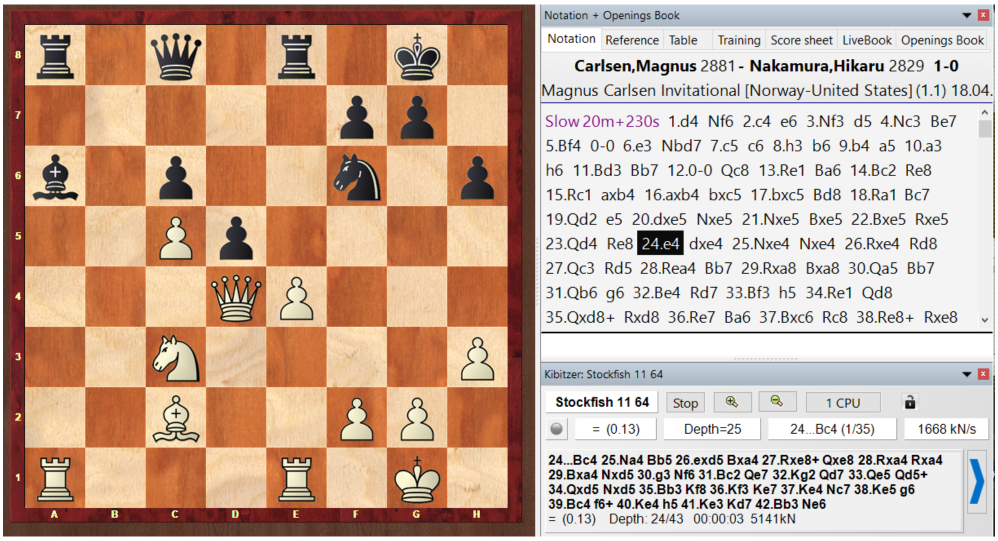
Zegners, Dainis, Uwe Sunde, and Anthony Strittmatter, "On the causes and consequences of deviations from rational behavior," *mimeo, LMU Munich*, 2020.

A Appendix

Table A.1 Descriptive statistics

	Number ind. moves	Mean	s.d.	Min	Max
Outcome variables					
Make Error	27,267	0.556	0.497	0	1
Error if Make error = 1	15,173	1.980	18.634	0.010	326.960
Control variables					
Complexity nodes	27,267	5,517,568	7,040,391	56	4.62×10^8
Elo score player	27,267	2,741.436	126.736	2,003	2,908
Difference in Elo score to opponent	27,267	2.400	176.278	-622	870
Number of games played before game g					
within the tournament	27,267	9.282	7.366	0	35
during a specific day	27,267	1.809	1.312	0	4
Remaining time before move (in min)	27,267	6.424	4.635	0.117	19.233

Note: The table shows the descriptive statistics based on the main estimation sample, i.e., excluding the first 15 moves of each game.



(a) Engine evaluation before move

(b) Engine evaluation after move

Figure A.1 Computation of error variable: An example

Note: The two consecutive positions above are taken from a game in our dataset. Before the black player made his move on move 24 of the game (upper panel), the chess engine evaluates the position with a pawn metric of +0.13 in whites favor, which corresponds to a disadvantage of -0.13 pawn units for the black player. The optimal move for the black player according to the chess engine is bishop to c4. However, the black player chose to play pawn takes e4 (lower panel). After this move, the pawn metric increases to +1.09 in whites favor, or -1.09 pawn units from the perspective of black. The error of black is computed as -0.13 - (-1.09) = 0.96. To compute the best move with a search depth of 25, the chess engine calculated 5,141,000 nodes (or 5,141 kilonodes) of the game tree in the position before the move, which corresponds to our measure of complexity.

Table A.2 Full estimation results

	Make Error	Ln Error if Make Error = 1
Online	0.021	0.168***
	(0.138)	(0.002)
ln(ComplexityNodes)	0.120***	0.370***
	(0.000)	(0.000)
ELO score player	0.000	-0.001**
	(0.836)	(0.019)
Difference in ELO score to the opponent	-0.000	-0.000
	(0.269)	(0.869)
Number of games played before game g		
within the tournament	-0.000	-0.004
	(0.871)	(0.219)
during a specific day	0.004	0.020
	(0.342)	(0.160)
ln(Remaining time before move in min)	-0.021***	-0.132***
	(0.001)	(0.000)
|Pawn metric of the position before the move|	0.002***	0.079***
	(0.000)	(0.000)
|Pawn metric of the position before the move| (squared)	-0.000***	-0.000***
	(0.002)	(0.000)
Constant	-1.254***	-4.921***
	(0.000)	(0.000)
Number ind. moves	27,267	15,173
R^2	0.178	0.191
Controls	YES	YES
Player FE	YES	YES
Move FE	YES	YES

Note: The table shows the full estimation results based on equation 2. Standard errors are clustered at the game level and p-values are reported in parenthesis. Section 2.2 describes the construction of the outcome variables. The opening phase of each game is excluded for each player ($m \leq 15$). *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$