Cognitive performance in remote work - evidence from professional chess

Cognitive performance in remote work

Steffen Künn¹, Christian Seel², and Dainis Zegners³,*

Abstract: During the COVID-19 pandemic, traditional (offline) chess tournaments were prohibited and instead held online. We exploit this unique setting to assess the impact of remote-work policies on the cognitive performance of individuals. Using the artificial intelligence embodied in a powerful chess engine to assess the quality of chess moves and associated errors, we find a statistically and economically significant decrease in performance when an individual competes remotely versus offline in a face-to-face setting. The effect size decreases over time, suggesting an adaptation to the new remote setting.

Keywords: remote work; productivity; chess

Classification: C72, J24, Z20

1 Introduction

Remote work (also known as telecommuting or working from home) has seen a steep increase during the COVID-19 pandemic. Surveys report the share of workers working remotely

*Correspondence address: zegners@rsm.nl, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands.
during the pandemic as 50% in the U.S. (Brynjolfsson et al., 2020) and 17% globally (Soares et al., 2021). Although the recent increase in remote work has been driven both by voluntary and mandated social distancing during the pandemic, it is arguably an acceleration of a broader trend toward more flexible work arrangements (Mas and Pallais, 2020), recent innovations in digital technologies (Bloom et al., 2021), and the rise of online labour markets (Agrawal et al., 2015). Dingel and Neiman (2020) estimates 37% of jobs in the U.S. could be done entirely remotely. Most firms will therefore face decisions on the scale and scope of allowing their employees to work remotely.

Given this increase in remote work, knowing how remote work affects workers’ productivity and which tasks are more suitable to being performed remotely is important. Yet, despite the large managerial relevance, the empirical evidence 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 remotely from changes in individual productivity. We contribute toward filling this gap by analysing the performance of professionals in one specific, well-defined cognitive task: playing chess.

During the COVID-19 pandemic, the current chess World Champion, Magnus Carlsen, initiated an online tournament series, the Magnus Carlsen Chess Tour. We analyse the performance of players who participated in these online tournaments and the performance of players participating in recent editions of the World Rapid Chess Championship as organized by the World Chess Federation in a traditional offline format. In particular, our main comparison is based on 20 elite chess players who competed both in the online and offline tournaments. We selected these tournaments because they were organized under comparable conditions.
conditions, in particular, giving players the same amount of thinking time per game, offering comparable prize funds, and implementing strict anti-cheating measures.

We base our performance benchmark on evaluating the moves played by the participants using a currently leading chess engine that significantly outperforms the best human players in terms of playing strength. We use the engine’s evaluation to construct a measure of individual performance that offers a high degree of objectivity and accuracy. Overall, we analyse 214,810 individual moves including 59,273 moves of those 20 players who participated in both the remote online and the traditional offline tournaments. Using a regression model with player fixed effects that allows us to estimate changes in within-player performance, we find the quality of play is significantly worse (at a statistical significance level of 5%) when the same player competed online versus offline. The adverse effect is particularly pronounced for the first two online tournaments, suggesting a partial adaptation to the remote setting in later tournaments.

A possible explanation for the drop in quality of play are weaker peer effects, which might induce higher concentration in a playing hall. Peer effects have been found to be important in various settings such as the workplace (Falk and Ichino, 2006; Mas and Moretti, 2009; Cornelissen et al., 2017), educational attainment (Zimmerman, 2003; Sacerdote, 2011), or sports competitions (Guryan et al., 2009; Brown, 2011; Hickman and Metz, 2018). Anecdotal evidence supports the important effects of peers and a playing venue different from home, because the World Champion and his coach rented a house as a playing site for later online tournaments to recreate some of the “tournament atmosphere”.

In principle, cognitive performance could also drop in response to factors related to the COVID-19 pandemic, such as uncertainty, anxiety, or income loss (Brodeur et al., 2021). On the other hand, Papageorge et al. (2021) document a larger burden for individuals with lower incomes, whereas we focus on wealthy elite-level chess players. Thus, we conduct a robustness analysis with a measure of pandemic-related regulations in a player’s home country as a control variable. The general pattern of our results - a strong decline in performance in the first tournaments - remains. As such, we find the explanation via peer effects more plausible.

Drawing general conclusions from one specific setting studying one particular type of worker is difficult, because the impact of remote work likely depends on the specific composition of tasks performed by a worker. Instead, our setting should be seen as a benchmark to study performance in a purely cognitive task. Cognitive tasks are important in many modern professional, managerial, technical, and creative occupations (Autor and Price, 2013), and cognitive skills have also been estimated to be especially important in occupations that can be conducted remotely (Malkov, 2020).

The observed decrease in the performance of chess players when competing remotely is likely to be smaller than the decrease in the performance of most other workers for several reasons. First, chess players are already used to remote work when preparing for their games with a computer, often working with a team of coaches who work remotely. They were even already familiar with playing chess online, albeit with shorter time controls and much lower incentives. Additionally, the players in our sample are relatively young (mostly between age 20 and 40) and wealthy, which might reduce potential frictions in adapting to the new situation. Finally, some players invited their coaches or worked from a remote location.

\footnote{For example, see \url{https://www.youtube.com/watch?v=gme35o8XvQw} (accessed on April 30, 2021).}
other than their homes to increase focus, a mitigation strategy that is not possible for many employees. On the other hand, once COVID-restrictions are lifted, only workers who feel comfortable adapting might select into remote work.

1.1 Relation to the Literature and Other Explanations

We contribute to the literature examining the impact of management practices, in our case, remote work, on productivity and firm performance (Bloom and Van Reenen, 2007, 2010; Bloom et al., 2012). In a meta-study, Gajendran and Harrison (2007) find no effect of remote work on self-reported performance elicited in unincentivised questionnaires and a positive effect of teleworking on supervisor-reported or archival records of performance. They conclude, “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, for example, Bailey and Kurland (2002) and Allen et al. (2015).

A seminal paper in the economics literature is Bloom et al. (2015), who examine the productivity of call-center workers in a randomised 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 find no statistically significant effects. In contrast to Bloom et al. (2015), we consider a highly specialized cognitively demanding task. Moreover, 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 remotely and are purely driven by task-level cognitive performance.

A few other recent studies are related. Using public sector data, Linos (2016) finds in a within-subject design that teleworking patent officers have 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 randomised into a treatment that allows for more flexible work arrangements in terms of hours worked and location. In a lab experiment, Dutcher (2012) finds a negative impact of conducting a dull work task online (typing numbers and letters on a computer keyboard, mimicking data entry) and a positive impact on conducting a creative task online (playing tic-tac-toe against a computer).

Other papers that look at the productivity of remote workers during the pandemic are Gibbs et al. (2021), who find a decrease in productivity of IT professionals when working from home, and Harrington and Emanuel (2021), who find an increase in productivity of call-center workers when moving to working from home during the pandemic. Beckmann (2016) (p.8) claims that for call-center employees, there is a “scope for productivity enhancements because employees working in large and noisy offices were easily distracted”. We do not think noise and distraction levels play a major role in our study, because they are low in offline tournaments and we are not aware of any distractions caught on camera during the online tournaments. In addition, players were highly incentivised to create a quiet working environment when competing remotely.
2 Data and Methods

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., 2021), 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 2018 - 2019 played offline in Saint Petersburg and Moscow, and from the Magnus Carlsen Chess Tour and its sequel, the Champions Chess Tour, both played online from April-November 2020 on the internet chess platform chess24.com. The raw data on the selected tournaments, including the exact documentation of all moves and thinking times, are downloaded from Chess24 (2021). Table 1 provides an overview of the tournaments contained in our dataset. All included tournaments are identical with respect to the time budget of 15 minutes per player to complete the game plus 10 seconds added to a player’s time budget for each move played. In contrast to shorter Blitz games (35-minute time budget per player), small differences in the time necessary to physically execute a move and press the clock relative to using a mouse to enter a move on a computer are unlikely.
to have a significant impact on the outcome in relatively longer rapid games. Finally, the majority (20 out 28) of players in the online tournaments also competed in at least one edition of the *World Rapid Chess Championships* in the years 2018 - 2019, enabling us to make within-player comparisons of performance for each of these 20 players.

Table 1: Overview of selected tournaments

<table>
<thead>
<tr>
<th>Tournament</th>
<th>Timespan</th>
<th>Prize Pool</th>
<th>First Prize</th>
<th>Registered Players</th>
<th>Estimation sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Offline tournaments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>World Rapid Chess Championships 2018</em></td>
<td>Dec. 26 - Dec. 28, 2018</td>
<td>$350k</td>
<td>$60k</td>
<td>206</td>
<td>1,265</td>
</tr>
<tr>
<td><em>World Rapid Chess Championships 2019</em></td>
<td>Dec. 26 - Dec. 28, 2019</td>
<td>$350k</td>
<td>$60k</td>
<td>207</td>
<td>1,434</td>
</tr>
<tr>
<td><strong>Online tournaments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Magnus Carlsen Chess Tour</em></td>
<td>Apr. 18 - May 3, 2020</td>
<td>$250k</td>
<td>$70k</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><em>Lindores Abbey Rapid Challenge (LARC)</em></td>
<td>May 19 - June 3, 2020</td>
<td>$150k</td>
<td>$45k</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td><em>Chessable Masters (CM)</em></td>
<td>June 20 - July 5, 2020</td>
<td>$150k</td>
<td>$45k</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td><em>Legends of Chess (LoC)</em></td>
<td>July 21 - Aug. 5, 2020</td>
<td>$150k</td>
<td>$45k</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td><em>Champions Chess Tour</em></td>
<td>Nov. 22 - Nov. 30, 2020</td>
<td>$100k</td>
<td>$30k</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>294</td>
<td>3,435</td>
</tr>
</tbody>
</table>

* The discrepancy between the number of registered players and players in our estimation sample is due to players not showing up for the tournament for example due to missed international flights. Games that are shorter than 16 moves have been dropped from the sample.

The *World Rapid Chess Championships* had an overall prize pool of $350,000 both in 2018 and 2019, with 35 prizes ranging from $60,000 to $1,500. Both tournaments had around 200 players participating, with many of the world’s elite among them. The tournament format was a 15 round Swiss tournament; that is, players with similar rankings in the tournament standing were paired against each other in each round, but the same opponents could only play each other once. The winner was the player with the highest score out of 15 games.\(^4\)

To prevent cheating, certified walk-through metal detectors at the entrance of the playing

\(^4\)In the case of a tie, a playoff with shorter Blitz games took place to determine the World Rapid Chess Champion. We disregard such games in our analysis.
hali enabled the tournament organisers to prevent players from bringing an electronic device with them.  

The prize pools in the online tournaments ranged between $100,000 and $250,000, with a prize span between $30,000-$70,000 for the winner and $1,500-$4,000 for the last place. These tournaments included 28 players (8-16 per tournament), each ranked in the top 100 of the official “FIDE World ranking” list for classical chess. These tournaments differed from other online tournaments in terms of the strict anti-cheating measures that included arbiters monitoring players via cameras at all times and standard automated cheating-detection systems in place. The tournaments were split into two phases, a league (either in one or two groups) with one mini-match of two to four games between each pair of players in the group and a knockout phase for the top players. In the knock-out stage (starting with a quarterfinal or semifinal depending on the size of the original field), players were paired in groups of two with seeding according to the results in the league stage and again played mini-matches of four games. The first to win a pre-specified number of mini-matches qualified for the next round.

In the empirical analysis, we include all games that were played in the World Rapid Championships 2018 and 2019 as well as the Magnus Carlsen Chess Tour and the follow-up Champions Chess Tour series until November 2020, except for the “Grand Final” of the “Magnus Carlsen Chess Tour,” because it only included four players who played a knockout

---

5 An electronic device could either be directly equipped with a chess engine or used to get information via a partner in the form of messages.
6 The automated system flags a player as being suspected of cheating if agreement with the optimal moves suggested by a chess engine is too large or if thinking-time patterns are indicative of cheating. Moreover, several commentators agreed that given their high standing in the world rankings, players would be very careful to avoid any suspicion of cheating, because this accusation would greatly damage their reputation, for example, [https://en.chessbase.com/post/magnus-carlsen-invitational-2020-preview](https://en.chessbase.com/post/magnus-carlsen-invitational-2020-preview) (accessed on June 10, 2020).
7 Details on the slight differences in procedures between the online tournaments can be found at [https://events.chess24.com/tour/](https://events.chess24.com/tour/) and [https://championschesstour.com/](https://championschesstour.com/) (both accessed on July 2, 2021).
tournament. 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 memorise as part of their preparation and training using engine analysis and opening books. In total, we observe 45,284 (169,526) moves played in 736 (2,699) games from the online (offline) tournaments.

2.2 Evaluation of Chess Moves

To estimate the effect of playing online on chess players’ performance, we evaluate each move in each game in our sample using the chess engine Stockfish 11. During the last decade, Stockfish has been consistently ranked first or near the top among chess engines, and it considerably outperforms every human player on off-the-shelf computer hardware in terms of ELO rating (Elo, 1978), the method used by the World Chess Federation to measure the strength of a player.

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 \(i\) (the last move \(M_g\) might only feature one individual move by the player who has the White pieces). For a given position in game \(g\) before individual move \(m_{ig}\), the chess engine computes an evaluation of the position in terms of the pawn metric \(P_{igm}\). Because 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.
advantage from the perspective of player $i$, with one unit indicating an advantage that is comparable to being one pawn up.\(^9\)

The pawn metric is computed by assuming both players play optimal moves; that is, the game proceeds along the optimal path computed by the chess engine.\(^10\) For each player $i$ in each game $g$ at each move $m_{ig}$, we compute two pawn metrics: $P_{igm}$ denoting the pawn metric before player $i$ makes his move and $P_{igm}$ denoting the pawn metric of the chess engine after player $i$ made his move. Using these two measures, we compute for each move an error defined as

\[
RawError_{igm} = P_{igm} - P_{igm},
\]

which reflects the change in the pawn metric after player $i$ has made his move $m_{ig}$. Intuitively, the $RawError_{igm}$ variable should be 0 if a move is deemed to be optimal and positive for erroneous moves. Yet, the evaluation function of the chess engine contains a small amount of randomness, which we account for with a random error term in our regression and in a separate robustness analysis in Section 3.\(^11\) We provide an example of the output of the chess engine and the computation of the error metric in Figure A1 of the online appendix.

\(^9\)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 into the pawn metric. See https://chess.fandom.com/wiki/Centipawn (accessed on June 16, 2020).

\(^10\)We restrict Stockfish 11 to a search depth of 25 moves ahead to economise on computing costs.

\(^11\)A seemingly attractive alternative would be to compute the evaluation of all feasible moves in a position and to define the error as the difference in evaluation between the optimal and the actual move played. However, this would increase the computation time to a prohibitively large degree as modern chess engines such as Stockfish 11 achieve their calculation speed by searching the game tree efficiently, eliminating non-optimal moves early on in the search process.
In addition, the chess engine displays the number of unique nodes of the game tree needed to reach a prespecified search depth, which we use for a measure of the (computational) complexity of the position.\textsuperscript{12}

### 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:

\[
\text{Error}_{igm} = \alpha + \delta \text{Online}_g + \beta \text{X}_{igm} + \eta_i + \gamma_m + \nu_{igm},
\]

where \(\text{Error}_{igm}\) is a continuous outcome measure for the severity of the error made in game \(g\) played by player \(i\) at move \(m\):

\[
\text{Error}_{igm} = \begin{cases} 
\ln(\text{RawError}_{igm} + 1) & \text{if } \text{RawError}_{igm} > 0 \\
0 & \text{if } \text{RawError}_{igm} \leq 0
\end{cases}
\]

which captures the measured error size for each move in our sample.\textsuperscript{13}

Our regression model includes a vector of game and move-varying controls \(\text{X}_{igm}\) that accounts for the following differences: First, we have move-specific controls that relate to the difficulty of the decision problem. More unbalanced positions are more difficult to handle, because often only one “correct” move exists, whereas more balanced positions tend to have multiple moves that do not change the evaluation drastically. To account for this, we control

\textsuperscript{12}The engine does not search all branches of the game tree, but selectively increases search depth for more promising alternatives. As such, the complexity measure indicates the engines difficulty in assessing the position. This complexity measure is also used in a chess setting by Alliot (2017), Strittmatter et al. (2020), Zegners et al. (2021), and Sunde et al. (2021).

\textsuperscript{13}We use \(\ln(\text{RawError}_{igm} + 1)\) instead of \(\ln(\text{RawError}_{igm})\) to ensure our combined measure is continuous at 0.
Cognitive performance in remote work

for the difficulty of the chess position by (i) the absolute value of the current pawn metric of the position before the player makes his move $P_{igm}$, as well as its squared term, and (ii) (computational) complexity as measured by the number of nodes the engine needed to evaluate to reach a prespecified search depth. We additionally include (iii) the remaining time budget of a player before each move as a measure of time pressure, which increases the difficulty of the decision problem. Second, we include a set of game-specific controls, including (iv) the current ELO rating of the player to move to account for changes in a player’s strength over time, as well as the ELO rating of the opponent to account for possible differences in performance and playing style against stronger or weaker opponents. Third, we control for tournament-specific factors, such as players potential fatigue, by taking into account (v) the number of games played before game $g$ within the tournament as well as during a specific tournament day. In addition to the controls, we include individual player $\eta_i$ and move number $\gamma_m$ fixed effects and an error term $V_{igm}$ clustered at the game level to allow for arbitrary correlation within each game. Table 2 shows the summary statistics of the variables used in the estimation.

The term $Online_g$ denotes the treatment indicator taking the value of 1 if game $g$ was played in an online tournament, and 0 otherwise. Our parameter of interest is denoted by $\delta$, which measures the difference in the quality of play between games conducted online and offline. We identify the parameter of interest by observing the same individuals $i$ playing

\[14\text{We use the official ELO ratings for rapid chess by the World Chess Federation; see https://ratings.fide.com/top_lists.php. Our effects are robust to including instead players' classical ELO ratings. These results are available from the authors upon request.}\]

\[15\text{We opted for clustering standard errors at the game level to account for correlation of performance between the two players competing in a game. In Table 3, we show our results are robust if we cluster at the player level.}\]
moves in the online and the offline tournaments.\footnote{Our data also include moves of players who did not compete online. Due to the individual player fixed effects, however, these players do not contribute toward estimating the main effect of playing online but allow for a more precise estimation of the parameters of the control variables.} Although we cannot make final statements concerning causality because of the absence of an experimental setting, the rich specification makes us confident that $\delta$ is likely to represent the causal effect of playing online (vs. offline) on the severity of the errors.

3 Results

Table 2 contains our main estimation results and shows the estimated coefficient $\hat{\delta}$ based on Eq. (2). 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 2. We find that playing online leads to a reduction in the quality of moves. The error variable as defined in Eq. (3) is, on average, by 1.7 units larger when playing online than when playing identical moves in an offline setting. This corresponds to a 1.7% increase of the measure (RawError+1) or an approximately 7.5% increase in the RawError.\footnote{The approximation adapts the approach of Halvorsen and Palmquist (1980) to the case of $\ln(y + 1)$ and computes the value at the mean of the dependent variables.} The effect is statistically significant at the 5% level.

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 (-0.0001308) is based on a regression without individual fixed effects,\footnote{Players’ ELO rating hardly changes over the limited observation period of three years. Therefore, we need to exploit the variation in the ELO rating across players to reliably estimate the impact of the players’ ELO score on the error.} indicating that if a player’s ELO rating increases by one point, the error variable as defined in Eq. (3) is reduced by 0.013 units.
Table 2: Main Results: Offline vs. online tournament setting on performance of chess players

<table>
<thead>
<tr>
<th>Outcome variable: Error</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Dummy</td>
<td>0.001</td>
<td>0.027***</td>
<td>0.017**</td>
</tr>
<tr>
<td></td>
<td>(0.921)</td>
<td>(0.000)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Number of moves</td>
<td>214.810</td>
<td>214.810</td>
<td>214.810</td>
</tr>
<tr>
<td>Player FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Move FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Move-specific</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Player-specific</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Tournament-specific</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

The table shows the estimated coefficient $\hat{\delta}$ based on Equation 2. P-values are reported in parenthesis and are based on clustered standard errors at the game level. Section 2.3 describes the construction of the outcome variable. The set of control variables includes:

(i) move-specific characteristics containing a measure representing the complexity of the position before the move was made, the remaining time before each move, and the absolute value of the current pawn metric of the position before the player makes his move $P_{ym}$ as well as its squared term,

(ii) player-specific controls including the current $ELO$ rating of the player as well as the difference in the $ELO$ rating to the opponent, and

(iii) tournament-specific factors describing the number of games played before game $g$ within the tournament as well as during a specific day. The opening phase of each game is excluded for each player ($m \leq 15$). Descriptive statistics of the included variables are shown in Table ??.

Full estimation results for the final specification (column 3) are shown in Table A1 in the online appendix.

\*: $p < 0.1$, \**: $p < 0.05$, \***: $p < 0.01$.

on average. Playing online increases the error variable, on average, by 1.7 units, which corresponds to a loss of 130 points of $ELO$ rating. The factual drop in playing strength, however, is likely to be lower, because our analysis excludes the opening stage of the game, which is less likely to be affected by the online setting. Moreover, our linear regression model might not account for smaller average error margins at the top of the $ELO$ distribution.

In addition to Table 2, we provide the treatment effect ($\hat{\delta}$ based on Eq. (2)) for each online tournament separately in Figure 1. As described in Section 2.1, the selected online tournaments took place consecutively between April and November 2020. The figure shows the negative effect of playing online on the quality of moves is strongest for the first (and
The adverse effect of playing online on the quality of moves decreases over time, possibly because players adapt to the remote online setting.\textsuperscript{19}

We further test the sensitivity of our results with respect to (i) sample restrictions, (ii) the definition of the outcome variable, and (iii) alternative estimation techniques and model specifications. Table 3 summarises the results of the sensitivity analysis. First, to improve

\textsuperscript{19}In an additional analysis, we include an interaction effect between the online dummy and a trend variable indicating the order of the online tournaments. The coefficient on the interaction effect is $-0.005$ with a p-value of 0.038, indicating a statistically significant decrease of the online effect. The full estimation results are available upon request.
the estimation of the control variables, our sample includes games and moves of players who have not competed online. Most of these players are professional chess players, but they are, on average, weaker than the online players. Therefore, we re-estimate the model, excluding offline games in which no player competed in an online tournament (column 2) and dropping all moves by players who only played offline (column 3). The table shows the magnitude of the effects is almost identical to the main results (column 1), which is as expected since the estimation using individual fixed effects exploits only within-individual variation. The statistical significance decreases for the online dummy, most likely because of the reduced sample sizes and thus lower precision in estimating the impact of the control variables. Second, we restrict the sample to players who were located within Europe (including Western Russia) during the online tournaments. We find similar results (see column 4), mitigating concerns that time-zone differences and resulting different starting times of games might affect players’ performance. Third, we exclude moves in positions that are evaluated as $|P_{igm}| > 2$, indicating one player already faces a significant (dis)advantage, potentially altering players’ behaviour. The overall effect of playing online is reduced (column 5) and not statistically significant at conventional levels (p-value of 0.159). This finding suggests the overall effect is sensitive to including errors in positions that are already relatively (dis)advantageous for a player. The overall picture of a particularly pronounced decrease in the quality of play in the first two tournaments remains. Fourth, we apply a more restrictive definition of the error, giving less weight to marginal (extreme) errors by setting the error to 0 (5) for moves being annotated as the best possible move or having a change in the pawn metric $\leq 0.1$ ($> 5$). The results turn out to be robust (columns 6 and 7), indicating the main results are not driven by marginal (extreme) or mechanical errors created by the randomness in the
evaluation of the chess engine. Fifth, we apply a censored regression (Tobit) model given that we artificially censor our dependent variable at 0.\textsuperscript{20} In general, the point estimates for the parameters of interest increase (column 8). Although the effect on the aggregate online dummy is not statistically significant at conventional levels (p-value of 0.125), the particular reduction in move quality during the first two tournaments remains statistically significant.

Sixth, we cluster the standard errors at the individual level, assuming an individual-specific pattern in playing behaviour. Given the low number of online players (N=20), we construct p-values based on wild bootstrap clusters as recommended by Cameron et al. (2008). The significance of the estimated effects increases (column 9). Seventh, to mitigate concerns that results are related to the pandemic, we add a control variable to the regression model to capture the severity of regulations implemented in a player’s home country during the tournament times (Source: Oxford COVID-19 Government Response Tracker, Hale et al., 2020).

Although the aggregate online dummy reduces in size and significance (p-value of 0.172), presumably because lockdowns occurred only during the online tournaments, the effect pattern on the separate tournament dummies remains almost identical relative to the main results (compare column 1 with 10).\textsuperscript{21} The lack of a relationship between quality of play and the severity of the pandemic is also supported by the sensitivity analysis that excludes non-European players (column 4), because different continents were in different phases of the pandemic throughout the online tournaments. Finally, Figure A2 in the online appendix shows the marginal effects of playing online, estimated separately for each of the 20 players.

\textsuperscript{20}From a conceptual point of view, we argue that the 0s represent a “true” observation (i.e., the player did not make an error) rather than censoring, as a player is highly unlikely to outperform the chess engine. Therefore, we prefer the OLS model as our main estimation strategy.

\textsuperscript{21}The coefficient on the lockdown variables is also not statistically significant in itself. Further, including a variable accounting for the number of COVID deaths within the week prior to a game in a player’s home country does not affect the main results and also is not statistically significant itself. These results are available from the authors upon request.
who competed both in the online and offline tournaments. We find no particularly large outliers, mitigating concerns that our main effect is driven by a particular player.

4 Discussion and Potential Explanations

We have compared the performance of professional chess players in traditional offline tournaments to online tournaments in which players participated remotely. Our results show a clear decrease in overall performance in the remote setting, which is particularly pronounced at the beginning of the time period when chess players had to switch to the new setting. Thus, players seem to have adapted to the new remote work setting. Anecdotal evidence even suggests conscious adaptations; for example, the World Champion’s team rented a house together for some events to create a “tournament atmosphere” and “normal routine,” which he was missing while playing from home.\(^{22}\) For the organisation of future chess events, our results suggest a newly created “Hybrid format,”\(^{23}\) in which only some players participate remotely, might lead to disadvantages for players who cannot attend the event in person due to travel restrictions or an inferior economic situation.

Our sample of young and tech-affine professionals who can choose their workplace and even invite co-workers to the same location seems particularly favourable to remote work. As such, we think the initial drop in cognitive performance and the adaptation time might be even more pronounced for most other workers. This effect could be partially offset by a selection of more adaptable workers into remote work.


Other key factors in cognitive performance and remote work are noise and distraction levels. Because these factors are highly dependent on the exact job and factors such as care responsibilities at home, they are not universally favourable or detrimental to remote work. In our setting, we expect noise-related factors to be of minor importance, but we cannot fully rule out additional (care) responsibilities due to lockdowns and competing from home.

Moreover, our study deals exclusively with individuals performing their task on their own. Although this focus allows us to cleanly identify effects that are purely due to differences in cognitive performance on an individual level, many work tasks in the modern economy are performed by teams. The impact of remote work on teams is an additional dimension that firms have to consider when deciding on their remote work policies.24

Acknowledgments

We thank Johannes Carow, Juan Palacios, and Anthony Strittmatter for providing helpful comments as well as seminar audiences at University of Zürich, Maastricht University, Rotterdam School of Management, ifo Institute Munich, the Baltic Economic Associations online seminar, the SCECR 2021 workshop, and the ChessTech 2021 conference. We are also grateful to three anonymous reviewers and the editor for their suggestions, which helped us to further improve the paper.

24 See, e.g., Battiston et al. (2021), van der Meulen et al. (2019) and Dutcher and Saral (2012) for studies on remote work in teams. DeFilippis et al. (2020) and Yang et al. (2021) examine the effects of working from home during the COVID-19 pandemic on teamwork.
5 Supplementary data

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.5674798.

Affiliations

1 Maastricht University, IZA and ROA
2 Maastricht University, Department of Microeconomics and Public Economics
3 Rotterdam School of Management, Erasmus University
## Table 3: Sensitivity Analysis

<table>
<thead>
<tr>
<th>Outcome variable: Error</th>
<th>Main results</th>
<th>Online players only</th>
<th>Sample definition</th>
<th>Outcome definition</th>
<th>Estimation and Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(see Table 2)</td>
<td>Games(^a)</td>
<td>Moves(^b)</td>
<td>errors(^c)</td>
<td>Tobit(^d) S.E. COVID19 restr. (^e)</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Online Dummy</td>
<td>0.017**</td>
<td>0.015</td>
<td>0.014</td>
<td>0.020***</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.081)</td>
<td>(0.115)</td>
<td>(0.009)</td>
<td>(0.159)</td>
</tr>
<tr>
<td>Magnus Carlsen Invitational</td>
<td>0.020**</td>
<td>0.037***</td>
<td>0.036***</td>
<td>0.047***</td>
<td>0.015**</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.000)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Lindores Abbey Rapid Challenge</td>
<td>0.022*</td>
<td>0.023**</td>
<td>0.022*</td>
<td>0.023*</td>
<td>0.012*</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.040)</td>
<td>(0.055)</td>
<td>(0.062)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Chessable Masters</td>
<td>0.001</td>
<td>-0.000</td>
<td>0.004</td>
<td>-0.006</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.914)</td>
<td>(0.952)</td>
<td>(0.971)</td>
<td>(0.705)</td>
<td>(0.339)</td>
</tr>
<tr>
<td>Legends of Chess</td>
<td>0.020*</td>
<td>0.016</td>
<td>0.015</td>
<td>0.019</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.159)</td>
<td>(0.178)</td>
<td>(0.123)</td>
<td>(0.762)</td>
</tr>
<tr>
<td>Skilling Open</td>
<td>0.013</td>
<td>0.030</td>
<td>0.009</td>
<td>0.016</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.166)</td>
<td>(0.344)</td>
<td>(0.395)</td>
<td>(0.130)</td>
<td>(0.149)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(N) Number of individual moves</th>
<th>214,810</th>
<th>70,306</th>
<th>59,273</th>
<th>196,491</th>
<th>147,754</th>
<th>214,810</th>
<th>214,810</th>
<th>214,810</th>
<th>214,810</th>
<th>214,810</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Move FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Move-specific Control variables</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Player-specific Control variables</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Tournament-specific Control variables</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

The table shows the estimated coefficient \(\delta\) based on Equation 2. \(p\)-values are reported in parenthesis and are based on clustered standard errors at the game level (except stated otherwise). Statistical significance is indicated by \(^*: p<0.1, \quad ^{**}: p<0.05, \quad ^{***}: p<0.01\). Section 2.2 describes the construction of the outcome variables.

\(a\) Excluding games where none of the players competed in the online tournaments.
\(b\) Excluding moves by players who never competed in the online tournaments.
\(c\) Excluding players who participated in the online tournaments from outside Europe or Russia. In total, we exclude 10 (out of 20) players from India, USA and China.
\(d\) Excluding moves in positions with a pawn metric \(|P_{igm}| > 2\).
\(e\) We redefine errors as zero for moves with marginal errors between zero and 0.1, or moves being the best possible as indicated by the chess engine.
\(f\) Extreme errors (\(\text{RawError} > 5\)) are set to 5.
\(g\) We apply a censored regression (Tobit) model given that we censor our dependent variable at zero.
\(h\) We cluster the standard errors at the individual level to allow for arbitrary correlation within individuals. Given the low number of online players \((N=20)\) potentially violating the large-sample assumptions, we construct \(p\)-values based on wild bootstrap clusters as recommended by Cameron et al. (2008), using the \texttt{boottest.ado} command in \texttt{STATA} (see Roodman et al., 2019).
\(i\) We add a control variable capturing the severity of regulations related to the COVID-19 pandemic as implemented at a player's home country during the tournament times (Source: Oxford COVID-19 Government Response Tracker, Hale et al., 2020). The index is measured at a daily basis and captures the following categories: 0 - No measures; 1 - recommend not leaving house; 2 - require not leaving house with exceptions for daily exercise, grocery shopping, and essential trips; 3 - Require not leaving house with minimal exceptions.
References


