# The Impact of Indoor Climate on Human Cognition: Evidence from Chess Tournaments*

**Steffen Künn**[†] **Juan Palacios**[‡]
**Nico Pestel**[§]

**Working Paper**
April 29, 2019

**Abstract**

This paper studies the impact of environmental quality on the performance of individuals undertaking cognitively demanding tasks under time pressure. We link measures of indoor air quality and thermal conditions to the performance of chess players at official tournaments where players face strong incentives to exert high effort. We use a state-of-the-art chess engine to detect erroneous moves and evaluate the quality of the move by comparing the quality of a player's actual moves with the best moves proposed by the chess computer. The results indicate air pollution (PM2.5) is the deterring factor hindering cognitive performance. We find that an increase of 10 $\mu g/m^3$ raises the probability of making an error by 1.5 percentage points, and increases the magnitude of the errors by 9.4%. The impact of pollution is exacerbated by time pressure. When players approach the time control of games, an increase of 10 $\mu g/m^3$, corresponding to about one standard deviation, increases the probability of making a meaningful error by 3.2 percentage points, and errors being 17.3% larger. Our results have important implications for high-skilled office workers, in particular, for those executing non-routine cognitive tasks whose share is steadily increasing in developed countries.

**Keywords: Indoor environmental quality, worker productivity, cognition, chess.**
**JEL codes: D91, I1, J24, Q50, Z20.**

*The authors thank Rafael Suchy and Sergej Bechtoldt for excellent research assistance in collecting the data. Moreover, we gratefully acknowledge the financial support by the Graduate School of Business and Economics at Maastricht University as well as the Institute of Labor Economics (IZA).

[†]Maastricht University and ROA, The Netherlands; IZA Bonn, Germany; s.kuenn@maastrichtuniversity.nl; Corresponding address: Maastricht University, School of Business and Economics, Department of Economics, Tongersestraat 53, 6211 LM Maastricht, The Netherlands. Tel: +31 433 882851, Fax: +31 433 882000.

[‡]Maastricht University, The Netherlands; IZA Bonn, Germany; j.palaciostemprano@maastrichtuniversity.nl

[§]IZA Bonn, Germany; pestel@iza.org

# 1 Introduction

The reduction of harmful effects of pollution and extreme temperatures on citizens' health and well-being is a major driver of environmental regulation in most countries around the world. Air pollution alone is estimated to be responsible for 9 million premature annual deaths, according to 2015 global estimates (Landrigan et al., 2018). Although the health-science literature has documented the detrimental effects of temperature or pollution on the respiratory and cardiovascular systems for decades, only recently has started to examine environmental effects on human brains and cognitive functioning (Underwood, 2017). The effects are substantial and might well have significant consequences for the human-capital formation of countries. Recent work documents significant drops in student performance when testing days coincide with high levels of air pollution (Ebenstein et al., 2016) or extremely high temperatures (Park, 2018).

The economic literature has devoted significant efforts to the estimation of the burden of hazardous environments on the labor supply. An increasing amount of studies are providing quasi-experimental estimates of substantial productivity losses associated with extreme temperatures or air pollution. So far, most of the existing evidence is based on routine manual occupations, such as agriculture workers or pear packers (Zivin and Neidell, 2012; Chang et al., 2016), where output is easy to measure, for example number of pears collected per hour. Our understanding of how these hazards affect the performance of workers in cognitive or analytical professions, where the value added of a worker tends to be much harder to quantify, is still limited. The initial studies in the field use measures such as quantity rates (e.g., number of calls handled per hour (Chang et al., 2019)), judges' decision time (Kahn and Li, 2019) or uptime (percent of time in a day that a trader is at his desk trading (Meyer and Pagel, 2017)) to estimate changes in the added value of a worker. However, the literature remains silent about how the final quality of the tasks or decisions undertaken by cognitive workers is affected by adverse environmental conditions.

This paper examines the causal impact of indoor environmental conditions on the cognitive performance of adults, using a pure measure of cognitive performance, assessing the changes in the quality of the tasks and not worker's availability to execute the task. [1] We use data on the performance of players in chess part of official tournaments in Germany, a cognitive demanding setting where subjects are under time pressure and face strong incentives to exert high effort. The computational nature of chess allows the construction of an objective outcome measure for

[1] Although most of the existing evidence linking pollution and temperature to worker productivity relies on outdoor measurements, the average American spends more than 90% of her time indoors. The exposure of building occupants to extreme temperatures or pollutants is likely to deviate substantially from outdoor levels. In addition to the intermediation of air conditioning or heating for temperatures, the U.S. Environmental Protection Agency (EPA) documents significant differences in certain pollutant concentrations between indoors and outdoors. Thus, we might well expect the environmental conditions of indoor workers to differ from the outdoor measurements.

cognitive performance by comparing the quality of a player's actual moves with the best move proposed by a chess computer program.

Our estimation sample contains performance data from 102 players making about 20,408 moves over a maximum of 14 matches (7 matches per tournament), with 44 players participating in both tournaments. The players differ in their skill levels ranging from beginners to advanced players. All players participating in the tournament have strong incentives to exert high effort and perform well throughout all tournament rounds, because the performance during the tournament counts for their chess rating score, which is a matter of prestige among chess players and has implications for future competitions. In addition, the tournament settings provide pecuniary incentives by awarding monetary prizes. Therefore, we identify the treatment effect by observing identical individuals playing multiple games under varying indoor environmental conditions.

The environmental measures used in this study come from three web-connected sensors located in the same room as the players studied, ensuring the accuracy of the exposure data. The sensors continuously measure the levels of CO2, PM2.5, and temperature for the entire study period. The games considered in our sample tournaments last about three hours on average, which is similar to the average exposure in epidemiological studies exploring the effect of CO2 or temperature on cognition (e.g. Satish et al., 2012). Thus, the exposure time of our subjects is sufficient to uncover an effect of the environmental conditions on their cognitive performance. The measured levels of CO2 range between 1,000 and over 2,250 ppm, temperature between 22 and 29 degrees Celsius (72 and 84 Fahrenheit), and fine particulate pollution between 12 and 58 $\mu g/m^3$.

Our identification strategy exploits the fact that players play games over several days, which enables us to examine the relationship between environmental conditions and cognitive performance across the same player's games. In a first step, we estimate the impact of environmental quality on the players' performance in a pooled sample. In our preferred specification, with player, year (tournament edition), round, and move fixed effects, and a set of control variables, we find no significant effects of temperature and CO2 on our error measures. For PM2.5, we find an increase in particulate pollution of 10 $\mu g/m^3$, about one standard deviation in the sample, leads to 1.5 percentage points increase in the probability of making a meaningful error, corresponding to an increase of 18.8% relative to the sample mean. In addition, we consider the magnitude of the error and find an increase of 10 $\mu g/m^3$ leads to 9.4% larger errors. The results are similar in magnitude to existing estimates within the literature.

In a second step, we split the sample by the number of moves within games (15–20, 21–30, 31–40 and >40 moves) to consider effect heterogeneity with respect to the impact of indoor environmental conditions on the performance of chess players with and without time pressure. The tournament rules set a time restriction of 110 minutes for the first 40 moves per player.

Therefore, decisions taken within the range of 31–40 moves can be assumed to be taken under relative time pressure, compared to other phases of the game. We find the negative impact of pollution on performance is exacerbated if decisions need to be taken under time pressure. Approaching the time control of games (31–40 within a game), an increase of 10 $\mu g/m^3$ raises the probability of making a meaningful error by 3.2 percentage points (equivalent to a 29.6% increase compared to the sample mean), with the errors being 17.3% larger. Again, we find no effects for temperature and CO2. A sensitivity analysis with respect to selective sample attrition as well as the implementation of falsification and specification tests support the robustness of the results.

Given that our results are based on levels of temperature and air quality that are within a moderate range, our results resemble the exposure of an average office worker in a Western economy. In addition, we argue our results are based on a cognitive task and therefore are likely to have strong implications for high-skilled office workers, in particular for those executing non-routine cognitive tasks. The roles of these tasks are gaining importance in developed labor markets and are represented in professional, managerial, technical, and creative occupations (Autor and Price, 2013).

The remainder of our paper is organized as follows. In the next section, we discuss the literature linking environmental conditions to human health and (cognitive) performance. In section 3 we provide a description of the game of chess and its use by the scientific literature to understand human behavior and performance. In this section, we also explain the construction of our performance measure and the estimation sample. In section 4, we present our empirical strategy. The results are presented in section 5 and robustness checks are shown in section 6. section 7 concludes

# 2 Literature

## 2.1 Environment and Health

For decades, the health science literature has linked exposure to high levels of air pollution or extreme temperatures with respiratory and cardiovascular mortality and morbidity. Air pollution alone is estimated to be responsible for 9 million premature annual deaths, according to 2015 global estimates (Landrigan et al., 2018). Similarly, a recent study based on 74 million deaths between 1985 and 2012 in 13 countries around the globe estimates that 7.7% of mortality was attributable to temperature exposures (Gasparrini et al., 2015). The health cost of hazardous environments goes beyond mortality. Quasi-experimental evidence shows how exposure to air pollution leads to an increase in direct medical costs, such as increases in hospitalizations and pharmaceutical expenses by individuals (Schlenker and Walker, 2016; Deschenes et al., 2017).

The health effects of air pollution have been the scope of the analysis of numerous studies over the past decades. A wealth of evidence suggests exposure to air pollution has detrimental consequences for human health, even at moderate levels. The inhalation of ozone or particulate matter has been associated with mortality and hospital admissions due to cardiopulmonary health problems (Brunekreef and Holgate, 2002). The respiratory system is the primary target of air pollutants. Epidemiological studies document associations between the presence of air pollutants and respiratory health morbidity, such as exacerbation of asthma or declines in lung function (for a review, see Rückerl et al., 2011). Cardiovascular systems are also vulnerable to airborne particles. For instance, exposure to high levels of ultra fine particles has been associated with the advent of ischemic heart disease or elevated blood pressure (Pope et al., 2004; Bhatnagar, 2006).

In addition, evidence on the mortality and morbidity attributable to hot and cold temperatures is increasing. The damaging effects of extreme temperatures generate important burdens for the cardiovascular and respiratory health systems. Epidemiological studies document significant links between temperatures and cardiovascular diseases such as atherosclerosis or pulmonary heart disease (Xiaofang et al., 2012). Global estimates indicate the existence of a significant burden of temperature on human health. Gasparrini et al. (2015) estimate the exposure to warm and (specially) cold temperatures is responsible for 7.71% of the total 74,225,200 deaths considered in the study. Although these estimates are mainly driven by exposure to sustained exposure to cold or warm temperatures, a recent wave of studies document substantial peaks in mortality associated with the advent of extremely warm and cold days (Barreca et al., 2016; Deschênes and Greenstone, 2011; Deschenes and Moretti, 2009).

Children and older adults tend to be the most vulnerable to air pollution or extreme temperatures. However, these hazards also have important implications for the working population; the economic literature has documented significant drops in the labor supply associated with both, temperatures and air pollution. The exposure to adverse environmental conditions has detrimental effects on the total time that workers supply. Hanna and Oliva (2015) shows, following the closure of a large refinery in Mexico City, an increase of 1.3 worker hours per week in the area surrounding the refinery. Similarly, (Aragón et al., 2017) document the negative impact of moderate increases in fine particles (PM2.5) on hours worked by people in Lima. Similarly, evidence from the US shows extremely warm days lead to a reduction in working hours, specially in industries with high exposure to outdoor climate conditions, such as the construction or forestry industry (Graff Zivin and Neidell, 2014). The authors find that on days with maximum temperatures above 30°C, workers in these industries reduce the time allocated to labor by one hour.

## 2.2 Environment and Cognition

Increasing evidence supports the hypothesis that environmental factors affect the human brain and cognition. The health science literature has provided convincing evidence on the detrimental links between poor air quality or thermal conditions and brain health and cognition. The economics literature suggest the cognitive impairment linked to these factors ultimately hampers academic achievement and human-capital formation.

A growing body of evidence in the area of epidemiology and toxicology suggests exposure to air pollution can harm the brain and hinder cognition via different channels. The inhalation of particles smaller than 200 nanometers can reach the brain, causing inflammatory reactions and ultimately impairing cognition (Underwood, 2017; Kumar, 2018). In adults, exposure to air pollution has been associated with depression, mood disorders, and ischemic strokes due to artery atherosclerosis or small vessel occlusion (Calderon-Garciduenas et al., 2015). Epidemiological studies have related long-term exposure to air pollution to the risk of dementia and cognitive decline in adult populations (Power et al., 2016). Finally, exposure to air pollutants such as carbon monoxide have been associated with a drop in the capacity of hemoglobin to transport oxygen, reducing oxygen availability to the brain and hindering concentration (Bernstein et al., 2004).

The consequences of pollution on cognition might be substantial even in the short term. Evidence from high-stake examinations in Israel shows the same student performs worse when the ambient levels of PM2.5 are higher during the exam day (Ebenstein et al., 2016). This finding is consistent with recent evidence linking the level of indoor PM10 to test scores of thousands of university students across multiple subjects in London. (Roth, 2018). Roth explores within-student variation in test performance under different levels of fine particles in the classrooms. The author finds significant substantial drops in grades when individuals take the exam in high-pollution days. The damaging effects of air pollution on human cognition seem to intensify with age. Zhang, Xin, Xi Chen and Zhang (2018) explore the effect of PM2.5 on survey-based test scores in verbal and numerical cognitive tasks in a representative sample of the Chinese population. The authors find the negative effects of pollution become more pronounced with age (specially for men).

Indoors, the main indicator used in the literature to measure the air quality in a room is the level of carbon dioxide (CO2). Humans are the main source of CO2, which is produced when breathing. The levels of CO2 in a room are mainly determined by the number of occupants and the ventilation rate in the room (rate at which the indoor air is exchanged with outdoor air). The inhalation of high levels of CO2 has been linked to fysiological and physiological symptoms such as dizziness, headache, or fatigue (Stankovic et al., 2016). Experimental evidence from the lab shows exposure to moderate levels of $CO_2$ impairs cognitive functioning of humans. In these studies, a group of individuals are asked to stay in a room where the levels of $CO_2$ have been artificially manipulated for several hours, and undertake a series of cognitive tests. The results of the studies indicate that even at moderate levels of $CO_2$ (e.g. 1,500 ppm), individuals performed significantly worse in cognitive tasks (Satish et al., 2012; Allen et al., 2016). In both studies, the experimentally induced high levels of $CO_2$ in the room mainly affect the cognitive domain of strategic thinking. A quasi-experimental study exploring the impact of a renovation program of ventilation systems in a sample of 65 US school buildings documents significant improvements in standardized test scores and passing rates (Stafford, 2015).

Finally, the literature also provides strong evidence on how exposure to extreme temperatures impairs cognitive function of humans. Recent experimental evidence based on functional magnetic resonance imaging (fMRI) shows alterations of brain blood flow upon exposure to heat stress (Liu et al., 2013). These alterations tend to impair individuals' ability to undertake complex tasks. In addition, the authors test a series of behavioral measures and find that high temperatures (50°C) impairs the executive function of individuals, but not the alerting and orienting functions. The results from a meta-analysis of the literature suggests heat stress affects only complex (cognitive) tasks, such as working-memory tasks, sustained attention, or tracking (Taylor et al., 2015). On the other hand, the cognitive studies in the area tend to show no significant effects of heat stress on simple cognitive tasks, such simple arithmetic tasks. Finally, a recent observational study using a series of cognitive tests suggest the reaction time of individuals increases with exposure to high temperatures. Cedeño Laurent et al. (2018) show that during a heat wave, participants that lived in air-conditioned houses have a 13% lower reaction time than their peers living in houses without air conditioning.

The impact of heat stress on cognition also translates into a reduction in test scores. Graff Zivin et al. (2018) analyze the effect of weather on cognitive performance of children using cognitive-assessment data from the National Longitudinal Survey of Youth. The authors find that daily changes in temperatures lead to substantial decreases in cognitive performance on math beyond 26 degrees Celsius, taking 21 degrees Celsius as the reference point. Given the voluntary character of the tests used as outcome in the study, the reduction in test scores is not purely the result of reduced cognitive performance, but a combination of a drop in cognitive abilities of participants and effort. Park (2018) links the scores of high-stake examinations in NYC to the ambient temperatures on the test day. Using the within student variation in exam temperatures across tests, Park finds that hot temperatures during exams results in reduced exam scores. Also, that this transitory shock in temperatures has long-term consequences, as reflected by the significant reduction in the likelihood of passing a subject at the end of the academic year. Recent evidence from a large sample of US exam scores shows high temperatures have been also associated with an impairment of learning (Goodman et al., 2018).

Cognitive ability has been considered an important factor in understanding how people perform and learn in strategic decision-making (for a review, see Rustichini, 2015). In a sample of 1,000 trainees, Burks et al. (2009) find that cognitive skills are a good predictor of strategic behavior. In particular, the authors find that individuals with high cognitive skills have significantly more accurate predictions of the behavior of the participants in a Prisioner's Dilemma game. In their seminal work, Gill and Prowse (2016) test in a large experiment whether cognitive ability influences individuals' ability to play a Nash equilibrium in a repeated game. The authors find that the individuals with higher cognitive skills perform better at the strategic game. In particular, in a p beauty contest game, more cognitively able participants choose strategies closer to the Nash equilibrium and learned faster than less cognitively able individuals.

## 2.3 Environment and Worker Performance

The harming effects of extreme weather conditions and pollution on labor supply go beyond drops in total working hours. Evidence from the manufacturing sector shows significant drops in the productivity of factories in periods with high pollution or high temperatures (Zhang, Xin, Xi Chen and Zhang, 2018; Fu et al., 2017). Similarly, a recent series of studies using daily productivity measures provides additional evidence on the harming effects of environmental conditions on workers' ability to exert high productivity levels. The studies focusing on manual routine occupations, such as agriculture workers or pear packers, show a reduction in performance of workers when exposed to high levels of pollution (Zivin and Neidell, 2012; Chang et al., 2016). These drops in productivity are also present in highly skilled, highly trained workers physical workers. Lichter et al. (2017) analyze the changes in the performance of professional soccer players associated with pollution in a sample of German first-league matches. The results show that variations in pollution across matches lead to changes in performance in soccer players. In particular, the number of passes per game that each player executes is reduced when the match takes place on a high-pollution day.

Physically demanding occupations are rely considerably on the respiratory and cardiovascular health of workers, which are heavily impaired by pollution and temperature. However, the results from a new series of studies suggest the productivity of office workers is not exempt from these hazards. Chang et al. (2016) explore how air pollution affects daily worker productivity of two call centers in China. The results indicate the number of daily calls handled by a worker decreases linearly with the level of local air pollution. The drops in daily productivity are not driven by the ability of the employees to handle the calls, but from an increase in the amount of time spent on breaks. Similarly, Meyer and Pagel (2017) link the daily trading activity of 103,000 private investors in Germany to contemporaneous levels of air pollution. The authors find that when investors are working on high-pollution days, they sit down less at their workplace, log in less often, and trade less in their brokerage accounts. Finally, a recent study examines the exposure to pollution and extreme temperates of 135,924 judges in 9.7 million criminal and civil cases (Kahn and Li, 2019). The authors show that exposure to high-pollution days leads to an increase in their total decision deliberation time period per case.

The evidence exploring the impact of extreme temperatures on worker performance relies mainly on call centers and lab studies where participants are asked to undertake several simulated office tasks (e.g., text processing). In a meta-review of the empirical literature, Seppänen et al. (2006); Seppänen and Fisk (2004) find the performance of participants tend to follow an inverted U-shape curve with the maximum at 21-24 degrees Celsius. Based on the meta-analysis of the studies, the authors estimate a 8.90% drop in individual performance associated with exposures to temperatures beyond 30 degrees Celsius The current evidence on how (indoor) environmental conditions (e.g., CO2 or temperature) affect the productivity of adult office workers is generally based on simulated office tasks that might well differ from real office settings. Individuals might undertake a series of behavioral responses (e.g., turn on a fan) to reduce the disutility produced by exerting high effort in high-temperature or highly polluted environments (Heal and Park, 2016). One of the likely reactions to adverse environments is the reduction of effort. As described above, evidence suggests a reduction in working hours or an increase in the time taken for breaks during the working days associated with pollution. Against this background, the estimates of the harming effects of extreme temperatures or poor air quality on performance from the lab studies where participants are not compensated by outcomes (but just by participating in the experiment) are likely to be a combination of a drop in pure cognitive performance and effort.

In sum, an increasing number of quasi-experimental studies provide evidence on the harming effects of pollution or temperature on worker productivity. Most of the current evidence relies on samples of manual routine jobs, such as agricultural or manufacturing workers. These occupations are usually physically demanding and therefore rely more heavily on the cardiovascular and respiratory systems than the tasks undertaken by office workers. The existing evidence on workers' productivity in cognitive professions uses piece rates or uptime as outcome variables (e.g., number of units produced or frequency and duration of times logged into a work station), but it is silent about the quality (or value) of the tasks undertaken by the subjects. The lab studies in the field complement the current field studies by looking at office-simulated tasks. However, the lack of incentives of participants makes differentiating between how much of the drop in performance comes from the drop in the ability of individuals to execute the task and how much from the drop in the effort in executing the task difficult.

This study deviates from the existing studies by exploring the effect of environmental conditions on cognitively demanding tasks in a setting where performance is remunerated, and thus the participants have clear incentives to exert high effort. In addition, this study is the first to investigate the impact on the quality of the produced outcome (cognitive performance) and not workers' availability to execute the task.

# 3 Chess Tournaments: Background and Data

In this paper, we use data from chess tournaments to study the impact of indoor environmental conditions on cognitive performance. Chess is a two-player strategic board game in which players alternately make moves with pieces on the chess board.[2] A player wins the game if (i) the player checkmates the opponent's king, (ii) the opponent resigns, or (iii) – in a game with time restrictions – the player runs out of time. In addition, the players can agree upon a draw at any time during the game.

Chess is a very complex, strategic, and computational activity, and has been heavily deployed by cognitive psychologists for investigating different strategic and cognitive aspects of human thinking, such as perception, memory, and problem solving (e.g. Charness, 1992). Burgoyne et al. (2016) provide empirical proof for the relationship between chess skills and general cognitive skills such as fluid reasoning, comprehension knowledge, short-term memory, and processing speed. In recent years, economists started using chess to analyze human behavior due its computational nature and the cognitive power of chess players (see, e.g., Palacios-Huerta and Volij, 2009; Gerdes and Gränsmark, 2010; Levitt et al., 2011; Backhus et al., 2016).

The data used in this paper come from two amateur chess tournaments in Germany. We received access to data on players' characteristics as well as the list of all moves of each individual tournament game. Throughout the tournaments, we measured indoor environmental conditions at the venue.

## 3.1 Tournament setup and chess rating score

The tournaments were organized by an amateur chess club in a major city in West Germany in May–June 2017 and April–May 2018 as the club's main event of the year.[3] Each tournament comprises seven rounds over an eight-week period with each round taking place on a Monday night starting at 6:00pm local time and lasting until the last game is over.[4] Figure A.1 in the Appendix illustrates the timing of the tournaments. Registration for the tournament was open to any amateur chess player on a first-come, first-served basis conditional on paying the participation fee of 30 euros. The total number of participants was limited to about 80 players per

[2]For details on the game of chess see the chess handbook as provided by the *World Chess Federation* (FIDE): `https://www.fide.com/fide/handbook.html?id=171view=article`.

[3] Further activities are participation in regional championship competitions, smaller-scale internal tournaments and regular training meetings.

[4] The weekly tournament rounds were paused for one week due to the public holidays Whit Monday (in 2017) and Easter Monday (in 2018).

tournament.[5] The tournament format follows the "Swiss system," a non-eliminating tournament format commonly applied in chess competitions. In each round, players gain one point for a win, 0.5 for a draw, and zero for a defeat. The winner of the tournament is the player with the highest aggregate points earned in all rounds. The assignment of fixtures is based on players' pre-tournament chess rating scores indicating their strength as well as their performance during the tournament.[6]

In general, chess rating scores are calculated based on the performance in games against other players. Winning (losing) a game results in an improvement (a decline) in the rating score, whereby the change in the rating score is larger in absolute terms for "unexpected" outcomes, for example, when a player with a much higher score than the opponent loses the game. The rating score applied for the assignment of fixtures in the tournaments is the German chess federation's rating score *DWZ* (*Deutsche Wertungszahl*).[7] This score is equivalent to the international *Elo* rating system as used by the world chess federation FIDE, also for assigning titles like "International Master" or "Grandmaster." We use the internationally acknowledged term *Elo* rating score instead of *DWZ* in the remainder of the paper.

After each tournament in our sample, all game outcomes are submitted to the chess federation for a recalculation of players' rating scores based on their results.[8] Hence, all players participating in the tournaments have an incentive to perform well throughout all tournament rounds in order to improve their rating score, which is a matter of prestige among chess players and which determines fixtures in future competitions. In addition, pecuniary incentives are offered. The winner of the tournament receives a cash prize of 400 euros. The participants ranked 2nd to 4th receive prizes of 300, 150, and 100 euros respectively, and extra prizes are awarded for the best-ranked players among the youth, the senior, and the female players (70 euros each), as well as for the best team (60 euros).

[5] Most participants are from the same city or from the surrounding region.

[6] Before the first round, all players are ranked based on their rating score. The ranking is then divided into the upper and lower half of the score distribution. In the first round, the highest-ranked player of the upper half (i.e., the player with the highest score overall) plays against the highest-ranked player of the bottom half (i.e., the player just below the median score) and so on. After round one, fixtures are assigned in the same way, but separately among the groups of players equal on points earned during the tournament. This implies that, by construction, the difference in rating scores between opponents is relatively high in the first round and typically becomes smaller in subsequent rounds because players with a higher score are more likely to win, especially when the difference is large.

[7] The *DWZ* rating system works as follows: Chess player $i$ is assigned a cardinal rating score $Z_{i,g}$ reflecting the player's strength before game $g$ against opponent $j$. The outcome of game $g$ determines the change in the score between games $g$ and $g+1$ according to the following formula: $Z_{i,g+1} = Z_{i,g} + \alpha_{i,g}[y_{i,g} - E(y_{i,g}|\Delta Z_{ij,g})]$, where the *actual* outcome for player $i$ in game $g$ is $y_{i,g} \in \{1, 0.5, 0\}$ for win, draw, or defeat, ereas the *expected* outcome is defined as $E(y_{i,g}|\Delta Z_{ij,g}) = \frac{1}{1+10^{(-\Delta Z_{ij,g}/400)}}$ based on the difference between players' scores, $\Delta Z_{ij,g} = Z_{i,g} - Z_{j,g}$, as well as a factor $\alpha_{i,g}$ depending on player $i$'s score level, experience, and age. See `https://www.schachbund.de/dwz.html` for details.

[8] The club has to pay a fee for the recalculation of participating players' scores, which is less expensive for the German *DWZ* score than for the international *Elo* score, which is why the organizers decided to "only" apply the *DWZ* score.

### 3.2 Move-performance measures

We measure the performance of players in each tournament round based on the quality of players' moves within the game. A chess game $g$ comprises $M_g$ moves, with two plies per move $m \in \{1, \ldots, M_g\}$, where the player with the white pieces moves first. For any given stage of the game, the relative (dis)advantage for each player is evaluated by the so-called *pawn metric* $C_{gm}$ based on the remaining pieces and their position on the board. Although it plays no formal role in the game, the pawn metric is useful to players and is essential to evaluate positions in chess software.[9] The sign of this metric indicates which player is in the better position (i.e., is more likely to win the game) with $C_{gm} > 0$ ($C_{gm} < 0$), indicating advantage for white (black). For example, a pawn metric of $-1$ is interpreted as the player with the black pieces having an advantage equivalent to one extra pawn on the board relative to the opponent.

For each tournament game, we have information on the evolution of the game based on players' hand-written notation (see Figure A.2 in the appendix for an example), which has been digitized by the tournament organizers.[10] We use a chess engine to assess the quality of each move in the tournaments. In theory, for each move, a particular move option optimizes the pawn metric given the situation on the chess board. Figuring out the best possible move is essentially a computational task for the human player. Therefore, we compare the pawn metric resulting from player $i$'s actual move $m$ in game $g$ with the metric that would result from the computer's optimally suggested move.[11] The pawn-metric difference between the human player and the computer can be viewed as an error:

$$Error_{igm} = |C_{igm}^{computer} - C_{igm}^{player}| \tag{1}$$

In the empirical analysis, we look at player-move specific errors as an outcome variable that may be affected by disadvantageous environmental conditions to which the players are exposed. We remove the first 14 moves of each game, which can be assumed to represent the opening game for which players usually have an established plan and are hence less affected by environmental conditions (Backhus et al., 2016). Furthermore, Expression (1) can take negative values when, at a given point in the game, the player makes a move that is evaluated to be better than the one proposed by the computer. Because we are mainly interested in the errors associated with the

[9]The metric values the remaining pieces on the board relative to a pawn, determining how valuable a piece is strategically. For example, knights and bishops are typically valued three times a pawn while the queen is valued at nine times a pawn. In addition, the value of a piece on the board differs depending on its position. See `https://chess.fandom.com/wiki/Centipawn` for details.

[10] Both players are obliged to document the evolution of moves and have to hand in the hand-written notation to the tournament organizer immediately after the game is completed. This notation is then submitted to the chess federation for the recalculation of players' rating scores.

[11]In this study, we use the chess engine Stockfish 9 64-bits with a current *Elo* rating score of 3548 (`http://ccrl.chessdom.com/ccrl/404/`). The highest *Elo* rating score by a human is 2882, achieved in 2014 by the current chess world champion Magnus Carlsen.

environmental conditions, and therefore the positive side of the error distribution, we redefine negative cases as zero (0.7 % of the sample). Panel A in Figure 1 displays the relationship between the average error per player and her *ELO* rating score, showing a clear negative relationship between the two. A statistically significant and negative correlation also exists between a player's *ELO* rating score and her mean error ($\rho = -0.54, p - value = 0.00$).

[Insert Figure 1 about here]

In addition to the continuous error measure, we explore the probability of an individual making a meaningful error based on the annotations of the chess engine. Chess engines are able to classify a certain move as a "meaningful error" based on the status of the game, the skill of the player, and the magnitude of the $Error_{igm}$. In particular, chess engines annotate a move $m$ as "meaningful error" if the engine considers move $m$ to be poor and should not be played weakening the chances of the player to consolidate her position or win the game. Given her skill level (*ELO* rating score), the player should be able to realize the move should not be played. The chess engine annotates two types of meaningful errors: (1) strategic mistakes and (2) tactical mistakes or blunders. The annotation of a move considered a strategic mistake describes a move that results in a loss of tempo or material for the player. These errors are considered strategic and not tactical. Blunders are severe errors that overlook a tactic from the opponent and usually result in an immediate loss in position, with a substantial drop in the chances of the player winning or drawing the game. The chess engine detects and annotates these errors. Panel B in Figure 1 displays the relationship between the average number of moves annotated as errors per player and the player's *ELO* rating score, showing a clear negative relationship between the two. The correlation between the average number of annotated meaningful errors per player (the sum of strategic mistakes and blunders) and her *ELO* rating score is -0.62 (p-value=0.00).

### 3.3 Time control

In each game, players face a time constraint (time control): Each player is allotted 90 minutes for the first 40 moves plus 30 seconds per completed move, resulting in a total time budget of 110 minutes for the first 40 moves. The time limit is allotted to each player individually and enforced by chess clocks. In each round, the tournament organizer announces the start for all games taking place in the same venue at the same time. If a player does not complete 40 moves within the time limit, he loses the game.

This measure gives each player a time budget to allocate to each move in the game, implying players are likely be under time pressure when they approach the 40th move and the time budget is reaching zero. To prevent losing the game altogether, a player then has to make move decisions more quickly, potentially within seconds, which makes them more prone to making lower-quality moves. Figure 2 shows the distribution of the total number of moves for all the games in our sample. The histogram shows peaks in the number of games finished around the move constraint (40 moves), suggesting that the imposed time constraint is binding, increasing the probability of ending a game right after the 40th move.

[Insert Figure 2 about here]

The assessment of quality at the move level allows for construction of our error measures throughout the game. In the empirical analysis, we exploit this feature of the tournament set-up to test whether the indoor environmental conditions during a game increases the effect of air quality or temperature on the probability of making errors when approaching the last move of the time control (move 40).

### 3.4 Measurement of indoor environmental conditions

During both editions of the tournament, the organizers granted us permission to measure indoor environmental conditions throughout all tournament rounds inside the venue, a large church community hall in a residential area. The players were informed that the measurement was being undertaken for scientific purposes. However, the players were not informed about the exact purpose of the study, namely, studying the effect of indoor environmental conditions on chess players' performance.[12]

Our measures of air quality (carbon dioxide, CO2, and fine particulate matter, PM2.5) and temperature were gathered from three real-time web-connected sensors located inside the tournament venue (see Figure A.3 in the appendix for an example).[13] The sensors measure the parameters of interest every minute and upload the measurements to a cloud server where the researchers can access the data in real time.

Figure 3 shows the distribution of the three parameters of interest over the seven rounds across the two editions of the tournament (2017 and 2018). The levels of CO2 range between 1,000 and over 2,250 ppm. These levels are above critical thresholds presented in the literature as detrimental for human cognition, for example, 1,000 or 1,500 ppm (Allen et al., 2016). The temperature levels during the tournament are between 22 and 29 degrees Celsius. Although these temperature levels are moderate, they are far from the temperature levels the literature considers as optimal for performance, namely, 21–24 degrees Celsius (Seppänen et al., 2006). Finally, the average level in our sample for PM2.5 is 25.9 $\mu g/m^3$, similar to the European target of $25 \mu g/m^3$ set by the European Environmental Agency (EEA, 2018).

[12] Just before the start of the first rounds, the main organizer of the tournaments informed all players about the presence of the sensors and that they should not be touched. In addition, we put signs next to each sensor explaining that the device was measuring indoor environmental conditions and should not be moved.

[13] We used two *Foobot* sensors and one *Netatmo* indoor sensor.

[Insert Figure 3 about here]

Note that important differences exists in the measurements of these parameters for the same rounds between the two years. In addition, no clear trend appears in the changes of the parameters between the years, but the changes in temperature or air quality between years are seemingly random. These differences are crucial for our estimation strategy, based on within-player and round variation of errors.

### 3.5 Descriptive statistics

Our data follow 102 players over a maximum of 14 matches. A total of 44 players participate in the two editions of the tournament. Table 1 shows summary statistics for player skills and demographic characteristics of the participants. Our sample is mainly composed of adult men who were, on averag,e 54 years old, with a wide range of levels of expertise. The least experienced player has only two official matches in her records and the most experienced player played 273 matches. The players also differ in their skills levels, according to the *Elo* rating score attached to their records. The *Elo* rating score of the most skilled player was more than twice as large as the *Elo* rating score of the least skilled player. In addition, Figure 4 shows the entire distribution of the *Elo* rating score of the players in the observed tournaments, and compares the scores with the official ranks within the chess association (FIDE). As the figure shows, we observe a wide range of skill levels ranging from beginners (novices) to advanced players (FIDE masters). In addition, the figure shows the *Elo* score of the chess engine *Stockfish* clearly dominating any human player.

[Insert Table 1 and Figure 4 about here]

Focusing on the game-specific characteristics (Panel B in Table 1), we can see that games in our sample last around three hours on average. This length is similar to the average exposure time in epidemiological studies exploring the effect of CO2 or temperature on cognition (e.g. Satish et al., 2012). In our study, the average game duration is sufficiently long to expect the exposure time of participants is sufficient to uncover an effect of the environmental conditions on their cognitive abilities. The average length of the games in our sample is around the 40-moves threshold (see Figure 2 for the full distribution of moves). About 20% of games finished in a draw.

Finally, the distribution of our outcome measures is shown in Panel C of Table 1. A total of 8% of the moves are annotated as meaningful errors. Moreover, 42% of the moves are considered suboptimal (positive error), with an average error rate of 1.43 pawns. Panel D in Table 1 shows the distribution of the indoor-environmental-quality variables within the estimation sample.

# 4 Empirical Model

Our goal is to estimate the effect of environmental conditions on the quality of the decisions undertaken by chess players. Our study setting has a number of features that allow us to identify the effect of environmental stressors on cognitive performance. First, players are executing the same (cognitive) tasks repeatedly in the same venue, the same day of the week, and at the same time of the day. In addition, the selection of opponents for each of the games is exogenously determined by the tournament organizer, following official rules in chess. Thus, participants have no control over the environmental conditions that they are exposed to during their games nor the opponents they play in a given round.

Second, we have objective measures of individual cognitive performance. In particular, we are able to evaluate each move in our sample of games. The chess engine is able to detect meaningful errors in the moves undertaken by the players. In addition, we build a continuous measure of the magnitude of the error. In fact, we can compute an optimum quality of the move, defined as the maximum (pawn) advantage that a player could reach if she would undertake the best possible move. In addition, we compute the advantage reached with the actual move of the player. The difference between the two is one of our main outcomes (see equation 1). The evaluation of the move quality is specific to the player's move and is not influenced in any way by the opponent.

Third, the high frequency of our outcome measures allows for the decomposition of the impact of environmental measures over different stages of the game. In particular, it allows us to test for differences in the magnitude of the impact as the time budget of players disappears over the course of the game.

Finally, all players in our sample face strong incentives to exert high effort and make optimal decisions, because the performance in each game of the tournaments counts for their chess rating score. Therefore, the incentive structure in our setting deviates from the structure in non-incentivized lab experiments or survey-based studies in which participants' payoffs are not determined by their performance in the proposed tasks. By contrast, our participants are highly motivated to perform to the best of their abilities.

We follow a fixed-effect strategy and estimate the following linear probability model:

$$Y_{ijtrm} = \alpha + \delta IEQ_{tr} + \beta X_{ijt} + \eta_i + \gamma_t + \lambda_r + \theta_m + V_{ijtrm} \quad (2)$$

where $Y_{ijtrm}$ is the outcome variable measured in a game between player $i$ and opponent $j$ at move $m$, round $r$, and year $t$. We consider two main outcome variables to capture the frequency and the magnitude of errors, namely, $MeaningfulError_{ijtrm}$ and $ln(Error_{ijtrm})$. $MeaningfulError_{ijtrm}$ is defined as a binary variable taking the value of 1 if move $m$, in round $r$, in year $t$, undertaken by player $i$ against opponent $j$ is annotated as a meaningful error. We consider meaningful errors those moves annotated by the chess engine as strategic mistakes and blunders (see section 3.2). We focus on annotated errors, instead of using the $Prob(error > 0)$, because not every positive error has a significant meaning for the game. For instance, some errors are minor without real consequences for the remainder of the game, or sometimes players create positive errors on purpose when they follow a risky strategy or try to force errors in the opponent. $ln(Error_{ijtrm})$ describes the natural magnitude of the error for individual $i$, playing against opponent $j$ in year $t$, round $r$ and move $m$ , describing the difference in the pawn metric between the computer's proposal and the player's move (see equation 1 for a detailed description of the variable).

We include a set of time-varying controls, describing the differences in skills between opponents in a given game, the points earned over the tournament by the player, and the initial advantage of the player before executing the move, pawn metric $C^{player}_{igm-1}$. We describe the differences in skills between the opponents with the variable $EloDiff_{ijt}$ that denotes the player-opponent difference in terms of the $ELO$ rating score to control for initial performance differences among the two players, measured at the beginning of the tournament. We include the variable $EloDiff_{ijt}$ as well as its squared term as controls. $\eta_i$, $\gamma_t$, $\lambda_r$, and $\theta_m$ are individual, year, round and move fixed effects, respectively. The parameter of interest is denoted by $\delta$, which measures the impact of prevailing indoor environmental quality $IEQ_{tr}$ on the outcome variable. In such a setting, the main identifying assumption is that pollution, temperature, and CO2 are assigned as good as randomly after including the rich set of fixed effects. Thus, we identify the parameter of interest by observing identical individuals playing against different opponents under varying indoor environmental conditions across tournament editions (years) of the same round of the tournament.

$IEQ_{tr}$ includes three available indoor environmental measures: (i) CO2 concentration, (ii) temperature, and (iii) fine particulate matter (PM2.5). All measures are included as the mean value of the prevailing conditions as measured during the second hour of the tournament rounds (N=14). Figure A.4 in the Appendix shows the distribution of the measures during the tournaments rounds. Whereas temperature and PM2.5 are relatively stable during the tournament rounds, CO2 concentration varies with the number of people in the room,namely, increasing (decreasing) at the start (end) of the tournament. Therefore, we decided to take the mean within the second hour of the tournament (as indicated by the dashed lines in Figure A.4) to avoid lower values at the beginning/end of the tournament polluting the measure.[14] Finally, the error term $V_{ijtrm}$ is clustered at the game level to allow for arbitrary correlation within the games in our sample.

[14] The replacement of our main regressors by the daily maximum values of these parameters does not change the results in sign and magnitude; see section 6.1.

# 5 Results

We present the results on the impact of environmental conditions on the performance of chess players in two stages: In a first step, the results based on the pooled sample are presented in section 5.1, where we estimate equation (2) using all moves in the games of the sample. In the second step, we split the sample into subsamples based on the status of a game, namely, the move number, in order to investigate effect heterogeneity with respect to time pressure. Players have a total of 110 minutes for the first 40 moves, inducing higher time pressure once they approach the 40th move than at the beginning of the match. The results for different moves levels are presented in section 5.2.

## 5.1 Pooled Estimation

Table 2 presents the estimated coefficients associated with environmental parameters in equation (2) using all moves in our sample. Panel A presents the estimation results using the probability of making a meaningful error as the outcome variable, and Panel B shows the results for the magnitude of the error ($ln(error)$). The columns in each of the panels display the estimates for a different set of fixed effects, starting with no fixed effects, and then stepwise including individual, year, round, and move-number fixed effects. All regressions include all environmental variables together with the set of control variables.

[Insert Table 2 about here]

Panel A in Table 2 shows the estimated coefficients $\hat{\delta}$ for the environmental parameters based on the regression as shown in equation (2). The outcome variable $MeaningfulError_{ijtrm}$ takes the value of 1 if the move is annotated as a meaningful error, and zero otherwise. With the most conservative specification (5) including the full set of fixed effects, we find no evidence for an effect of temperature or the concentration of CO2 in the room. The results indicate only the level of PM2.5 affects the probability of making a meaningful error. The significance and magnitude of the estimate even increases with the inclusion of additional fixed effects. The results of our main specification (5) indicate a 10 $\mu g/m^3$ increase in PM2.5 raises the probability of a player making a meaningful error by 1.5 percentage points in a given move of a game. This effect is equivalent to an 18.8% increase given the average probability of making a meaningful error in our sample of 8.0% (see Panel C in Table 1).

In Panel B in Table 2, we present the analysis of the magnitude of those errors, based on the estimation results of equation (2). The results are similar to Panel A. Although we do not find any significant effects for CO2 and temperature, the results show a significant impact of fine particles (PM2.5) on the magnitude of the error. For our main specification (5), we find that a 10 $\mu g/m^3$ increase in PM2.5 leads to a 9.4% increase in our error measure.

### 5.2 Effect Heterogeneity with Respect to Time Pressure

The time-control regulations of the tournament rules induce time pressure , requiring players to make the first 40 moves within 110 minutes of the game; otherwise, they lose the game. In this section, we estimate equation (2) for four different subsamples of move intervals within games, namely, 15–20 (24% of the sample), 21–30 (34%), 31–40 (22%), and >40 moves (20%). Decisions taken within the range of 31–40 moves can be assumed to be taken under relative time pressure, compared to the other categories given the low expected time left to execute the required 40 moves to stay in the game. In our sample, 44.4% percent of the games last more than 40 moves.

[Insert Figure 5 about here]

Figure 5 shows the estimated parameters with respect to the probability of making a meaningful error (Panel A) and the magnitude of the error (Panel B). All regressions contain individual, year, round, and move fixed effects, all environmental measures, and the full set of control variables: (i) the difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely, the pawn metric of the previous move by the opponent ($C_{jtrm-1}^{opponent}$). The dots represent point estimates and the black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level.

First, we focus on the results concerning the effect of environmental conditions on the probability of making a meaningful error (Panel A in Figure 5). In line with the results of the pooled regression, the results indicate CO2 and temperature have no effect on the probability of making a meaningful error at any stage of the game. However, we detect a clear pattern for the case of PM2.5. The estimated coefficients increase in size and significance the closer the game gets to the 40th move. This finding suggests that the effects as displayed in Table 2 are entirely driven by the moves close to move 40, when the time control takes place. Focusing on the move category 31–40, we find a 10 $\mu g/m^3$ increase in the levels of PM2.5 in the room leads to an increase in the probability of making a meaningful error by 3.2 percentage points. This effect is equivalent to a 29.6% increase given the average probability of making a meaningful error in our sample (10.8% for moves in this range).

Second, we focus on the impact of environmental conditions on the magnitude of the error. Panel B in Figure 5 shows the estimated coefficient $\delta$ of equation 2 using $ln(error)$ as the outcome variable. Again, we observe no evidence of a detrimental effect of CO2 concentration or temperature on the magnitude of the error. However, the estimates associated with PM2.5 show a positive and significant impact immediately before the time control at move 40. In earlier phases of the game before move 30, the existing variation in environmental conditions in our sample does not yield any impact on the performance of the chess players. In particular, we find that when games are in the move interval between 30 and 40, an increase in 10 $\mu g/m^3$ is associated with a 17.3% increase of the error.

In sum, we find the negative impact of fine particles on the performance of chess players is exacerbated by time pressure. The level of particles affect both, the probability of making meaningful errors during the game and the magnitude of errors. In our preferred specification, including the full set of fixed effects, the results from the pooled regression indicates an increase in 10 $\mu g/m^3$ (similar to one standard deviation in our sample) leads to an increase of 1.5 percentage points in the probability of making a meaningful error and 9.4% larger errors (relative to the average error). When we estimate the parameters for the move interval before the time control (move 30-40), the estimated coefficients double.

### 5.3 Effect Size in Context

Previous studies find a negative effect of CO2 on the cognitive performance of adults (e.g. Allen et al., 2016); however, the level at which CO2 impairs cognitive performance and the exact mechanisms for cognitive impairments remains unclear. In a lab experiment, Allen et al. (2016) shows that levels beyond 1,500 ppm have a detrimental effect on the performance of 24 adults in a simulated management task, using 500 ppm as a baseline. Zhang et al. (2015) reduce the air supply in the chamber to let subjects be exposed to 3,000 ppm of CO2. The authors find a cognitive impairment in the subjects at 3,000 ppm. The distribution of values of CO2 observed in our study differs from the distributions in lab experiments. Our baseline ($minCO2 = 1,179$ ppm) is twice the 500 ppm value commonly used in the literature as the reference CO2 level. We find no evidence that higher levels of CO2 are correlated with a higher presence of errors or the magnitude of the errors within the range of values considered in the analysis.

A number of studies show significant drops in the cognitive performance of humans under heat stress. In their meta-review of lab studies, Seppänen et al. (2006) find an average loss in cognitive performance of workers beyond 24 degrees Celsius - relative to the temperatures between 22 and 24 degrees Celsius. In their field study, Graff Zivin et al. (2018) find a significant drop in cognitive performance of subjects taking math tests while temperatures are above 26 degrees Celsius, using 22 degrees Celsius as the reference category. Our study spans temperatures between 22 and 29 degrees Celsius (72 and 84 Fahrenheit), with 40 percent of the rounds with average temperatures beyond 24 degrees Celsius and 27% of days beyond 26 degrees Celsius. We find no effects of temperature in a joint regression with CO2 and pollution (measured at the exact tournament time) on the quality of decisions of highly incentivized subjects.

Evidence on the impact of air pollution on cognitive performance of adults is increasing. Ebenstein et al. (2016) find that a 10-unit increase in daily PM2.5 (AQI) leads to an increase of 2 percentage points in the probability of failing a high-stakes exam. In our pooled sample, we find comparable effects with 10 $\mu g/m^3$ increase resulting in a 1.5 percentage points increase in the probability of making a meaningful error. Importantly, when looking at the move interval before the time control, we find the impact of PM2.5 doubles. An increase of 10 $\mu g/m^3$ in PM2.5 leads to a 3.2-percentage points increase in the probability of making meaningful errors. When looking at continuous variables of performance, we see heterogeneity in the elasticities of pollution on performance.[15] Among manual workers, the highest elasticity is 0.260, estimated in a US sample of agriculture workers (Zivin and Neidell, 2012). For China, Kahn and Li (2019) estimate the elasticity of PM2.5 in a sample of highly skilled public workers, finding elasticities between 0.179 and 0.243. In our pooled sample, we find a 0.267 elasticity associated with PM2.5. When restricting the sample to the move interval before the time control, we observe that the elasticity increases to 0.484, suggesting the effect of PM2.5 on cognitive performance is exacerbated under time pressure.

In sum, we find no impact of CO2 and temperature during the tournament rounds on the quality of tasks of our subjects. The estimated impact of PM2.5 in the full sample of moves suggests the existence of significant impairments of cognition, at a magnitude similar to the estimates of the literature. The estimates double in the sample of moves just before the time control. This observation suggests that when the time available for the execution of the cognitive tasks is limited, the impact of pollution increases substantially.

# 6 Sensitivity Analysis

In this section, we present a number of sensitivity tests to check the robustness of our significant results on pollution (PM2.5).[16] In particular, we reestimate the linear probability model as shown in equation (2), introducing the following modifications: (i) We use the daily maximum instead of mean value of the environmental parameters during the tournament rounds. (ii) We restrict the sample by removing games with less than 40 moves, and (iii) use data on outdoor pollution PM10 and ozone stemming from the closest air-quality stations. (iv) Finally, we provide a falsification test by additionally including pollution measurements the day before and after the tournaments rounds.

Figures 6 and 7 summarizes the findings with respect to sensitivity checks (i) – (iii) and show the estimated coefficients on the pollution parameter. Figure 8 shows the results of the falsification test. All specifications include the CO2 and temperature levels, the full set of fixed effects, and control variables as regressors.

[Insert Figure 6, Figure 7 and Figure 8 about here]

[15] See Kahn and Li (2019) for an excellent overview of the elasticities found in previous studies.

[16] We provide the results of the sensitivity analysis for temperature and CO2 in Figures A.5 and A.6 in the appendix, but refrain from discussing them here because we do not find any significant effects on these measures in the main analysis.

### 6.1 Maximum Values

We first test the sensitivity of the results with respect to measurement of the environmental conditions by using the maximum instead of the mean value of the air-quality measures and temperature as the treatment. Panel A in Figure 6 presents the results, which are consistent with our main estimates. The coefficients associated with PM2.5 remain significant and of a similar magnitude to those presented in the results section (see Figure 5).

### 6.2 Attrition

In our sample, a number of games do not get to the 40th move, when the time control takes place. Those games are likely to display differences in the number of errors in the earlier stages of the games that might lead to the early defeat of one of the players. These games might mislead our interpretation of the results, which might well be driven by those games finishing before the 40th move, and not by the time pressure induced by the time control per se. In this subsection, we present the estimation results restricting our sample to those games that reach the move 40.

Panel B in Figure 6 presents the estimation results of the main equations for the sample of games lasting at least 40 moves. The results suggest the main findings from section 5 are not driven by the games that finish before the time control is implemented. The estimates associated with PM2.5 do not change and even slightly increase in magnitude. In the move interval between 30 and 40 moves, a 10 $\mu g/m^3$ increase in the levels of PM2.5 in the room is associated with a 4%-points increase in the probability of making a meaningful error and 18.9% larger errors, compared to 3.2 percentage points and 17% in our unrestricted sample (see section 5.2).

### 6.3 Outdoor values

The existing studies in the field of environmental economics predominantly rely on outdoor measures of the environment (except for Roth (2018)). Thus, the existing studies tend to use data from weather stations (e.g., Park, 2018) or local air-quality stations (e.g., Ebenstein et al., 2016) to measure the exposure of individuals to certain temperatures or air pollution. In this subsection, we follow the traditional approach in the literature and replace our main regressors with outdoor measures. In particular, we replace the temperature and pollution treatments with the corresponding measures retrieved from an air quality and a weather station close to the tournament venue (about 3.8 kilometers). The outdoor measures are measured during the same time interval as the indoor measures, namely, during the second hour of the tournament rounds. However, for pollution, we have to rely on PM10 because PM2.5 is not available for the outdoor measurement.[17]

[17]Unfortunately, the outdoor measurement of PM2.5 is only available as the daily mean for every second day, so we decided to rely on the PM10 measurement instead.

Panel C in Figure 6 shows the results when we use the outdoor measure of PM10 instead of the indoor measure of PM2.5 as the treatment. We find an identical pattern for the coefficients on outdoor PM10, compared to our main results using indoor PM2.5 (see Figure 5). Within the category 31-40 moves, the magnitude and significance of the effects do not change. This finding is mostly attributable by the high correlation between the two pollution measures of 0.76 in our sample.

Finally, we test whether the effect is due to general pollution or is specific to PM2.5. We include the average level of ozone in the area during the tournament rounds in the main empirical model, together with PM2.5 and the rest of the environmental measures (equation (2)). Figure 7 shows the estimated coefficients associated with outdoor levels of PM10 and ozone. Although the coefficient associated with PM10 remains unchanged, ozone never has a significant effect in our sample. This finding supports the hypothesis that the estimated impacts of air pollution are mainly driven by the level of particles.

### 6.4 Falsification test

Our analysis so far has focused on the effects at the time of the tournament rounds. In this subsection, we present the results of a falsification test in which we estimate the relationship between the error measures and average pollution at times other than during the actual tournament rounds. In particular, we estimate a modified version of equation 2, from the pollution levels on days leading up to and following the tournament round.

We generate this mis-assigned pollution using the levels of PM10 corresponding to the second hour of the tournament rounds (7:00pm-8:00pm) in the two preceding ($t-2$ and $t-1$) and two following days ($t+1$ and $t+2$). [18] In addition, we include the pollution levels in the early morning (6:00am-9:00am) of the same day of the tournament round.

Figure 8 shows the results of seven separate regressions (including the pollution during the time of the tournament round). As anticipated, the observed positive relationship between the level of pollution and the error measures is strongest when we use the PM10 at the exact time of the tournament. The rest of the coefficients are not significantly different from zero. This finding is supportive evidence that our results on the probability and magnitude of errors are driven by the transitory effect of pollution, rather than by other explanations. The lack of effects of the lag PM2.5 indicates an absence of lagged health channels driving our performance measures. The absence of an effect for lead pollution offers further confirmation that our results are not driven by unobserved confounding factors.

[18] Given the lack of indoor measurements on the days before and after the tournament rounds, we rely on PM10 levels from the same air-quality station used in section 6.3 (3.8 kilometers away from the tournament venue).

# 7 Conclusion

In this paper, we investigate the impact of environmental conditions on human cognition by examining the performance of chess players at tournaments under different levels of air quality and temperature. Chess requires players to use their cognitive skills intensively and to decide strategically. Due to the computational nature of the game of chess, the cognitive performance of players can be measured very objectively by comparing the quality of a player's actual moves with those moves proposed by chess computer. In addition, chess players at tournament have a strong intrinsic as well as extrinsic motivation to exert high effort. By using this setting, we contribute to the existing literature on the effects of environmental conditions on human productivity, which so far have relied on using simulated office tasks in lab settings, and field studies focusing on routine manual occupations or workers' availability to execute tasks (or uptime) in non-routine cognitive occupations.

In addition, most studies are based on outdoor measurements of the environment that are likely to deviate from the actual environmental conditions (office) workers are exposed to during the working day. In our study, we were able to install measurement sensors recording the indoor environmental quality (CO2 concentration, temperature, and PM2.5) to which the players were exposed during the tournaments.

Our study is based on detailed move-level information collected at two chess tournaments in Germany. In total, we observe 102 players making 20,408 moves over a maximum of 14 matches (7 matches per tournament). Based on move-level information, we calculate our main outcome variable – the move-specific error rate – as the difference between the quality score of the actual move and the "optimal" move as proposed by a chess engine. To estimate the effect of indoor environmental conditions on the players' performance, we regress the error rate on the environmental conditions in the tournament round as well as individual, year, round, and move fixed effects. Further, we control for the difference in initial skill levels between the player and her opponent as measured by the *ELO* rating score, the number of points achieved during the tournament, and the actual status of the game before the move.

The results consistently indicate pollution harms the players' performance in cognitive tasks, whereas we find no effects for temperature and CO2 concentration. The estimation results show a 10 $\mu g/m^3$ increase in PM2.5 results in a 1.5-percentage-point increase in the probability of a player making a meaningful error, with 9.4% larger errors. The results on pollution are similar in magnitude to existing estimates within the literature. However, the effects double if decisions are taken under time pressure. We identify the different phases of the game by exploiting a tournament rule stating that the first 40 moves have to be completed within a total time limit of 110 minutes. We find a clear pattern showing the performance of players becomes even more sensitive to pollution when approaching move 40, when the time control takes place and the time budget is at its minimum. For the closest move category, 31-40 moves, we find a 10 $\mu g/m^3$ increase in the levels of PM2.5 in the room leads to a 3.2-percentage-point increase in the probability of making a meaningful error, and 17.3% larger errors.

Given that our measures of indoor environmental conditions are within a moderate range, resembling normal conditions humans are usually exposed to during their daily life, we argue that our findings can be extrapolated to different setups where individuals are required to make complex decisions or execute cognitive tasks under time pressure. For the labor market, given the type of cognitive task chess players have to perform (and which we actually measure with our outcome variable), our results likely have strong implications for the productivity of high-skilled office workers, in particular, for those executing non-routine cognitive tasks requiring problem-solving skills. Due to the technological change, the role of these tasks is steadily rising in developed labor markets and is represented in professional, managerial, technical, and creative occupations (Autor and Price, 2013).

## References

Allen, J. G., P. MacNaughton, U. Satish, S. Santanam, J. Vallarino, and J. D. Spengler (2016): "Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: A controlled exposure study of green and conventional office environments," Environmental Health Perspectives, 124, 805–812.

Aragón, F. M., J. Jose, and P. Oliva (2017): "Particulate matter and labor supply : The role of caregiving," Journal of Environmental Economics and Management, 86, 295–309.

Autor, D. and B. Price (2013): "The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003)," MIT Working Paper.

Backhus, P., M. Cubel, M. Guid, S. Sánchez-Pages, and E. Lopez Manas (2016): "Gender, Competition and Performance: Evidence from Real Tournaments," IEB Working Paper.

Barreca, A., K. Clay, O. Deschênes, M. Greenstone, and J. S. Shapiro (2016): "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the 20th Century," Journal of Political Economy, 124, 105–159.

Bernstein, J. A., N. Alexis, C. Barnes, I. L. Bernstein, J. A. Bernstein, A. Nel, D. Peden, D. Diaz-Sanchez, S. M. Tarlo, and P. B. Williams (2004): "Health effects of air pollution," Journal of Allergy and Clinical Immunology, 114, 1116–1123.

Bhatnagar, A. (2006): "Environmental cardiology: Studying mechanistic links between pollution and heart disease," Circulation Research, 99, 692–705.

Brunekreef, B. and S. T. Holgate (2002): "Air pollution and health," Lancet, 360, 1233–1242.

Burgoyne, A. P., G. Sala, F. Gobet, B. N. Macnamara, G. Campitelli, and D. Z. Hambrick (2016): "The relationship between cognitive ability and chess skill: A comprehensive meta-analysis," Intelligence, 59, 72–83.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2009): "Cognitive skills affect economic preferences, strategic behavior, and job attachment." Proceedings of the National Academy of Sciences of the United States of America, 106, 7745–50.

Calderon-Garciduenas, L., A. Calderon-Garciduenas, R. Torres-Jardon, J. Avila-Ramirez, R. J. Kulesza, and A. D. Angiulli (2015): "Air pollution and your brain: what do you need to know right now," Primary Health Care Research and Development, 16, 329–345.

Cedeño Laurent, J. G., A. Williams, Y. Oulhote, A. Zanobetti, J. G. Allen, and J. D. Spengler (2018): "Reduced cognitive function during a heat wave among residents of non-air-conditioned buildings: An observational study of young adults in the summer of 2016," PLoS Medicine, 15, 1–20.

Chang, T., J. G. Zivin, T. Gross, and M. Neidell (2016): "Particulate pollution and the productivity of pear packers," American Economic Journal: Economic Policy, 8, 141–169.

Chang, T. Y., J. Graff Zivin, T. Gross, and M. Neidell (2019): "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," American Economic Journal: Applied Economics, 11, 151–172.

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

Deschênes, O. and M. Greenstone (2011): "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US," American Economic Journal: Applied Economics, 3, 152–185.

Deschenes, O., M. Greenstone, and J. S. Shapiro (2017): "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program and Ozone Reductions," American Economic Review, 107, 2958–2989.

Deschenes, O. and E. Moretti (2009): "Extreme weather events, mortality, and migration," The Review of Economics and Statistics, 91, 659–681.

Ebenstein, A., V. Lavy, and S. Roth (2016): "The Long Run Economic Consequences of High-Stakes Examinations : Evidence from Transitory Variation in Pollution," American Economic Journal: Applied Economics, 8, 36–65.

EEA (2018): "Air quality in Europe," Tech. rep.

Fu, S., V. B. Viard, and P. Zhang (2017): "Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China," .

Gasparrini, A., Y. Guo, M. Hashizume, E. Lavigne, A. Zanobetti, J. Schwartz, A. Tobias, S. Tong, J. Rocklöv, S. Paulo, and S. Paulo (2015): "Mortality risk attributable to high and low ambient temperature : a multicountry observational study," Lancet, 386, 369–375.

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

Gill, D. and V. Prowse (2016): "Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis," Journal of Political Economy, 124, 1619–1676.

Goodman, J., M. Hurwitz, J. Park, and J. Smith (2018): "Heat and Learning," NBER Working Paper.

Graff Zivin, J., S. M. Hsiang, and M. Neidell (2018): "Temperature and Human Capital in the Short and Long Run," Journal of the Association of Environmental and Resource Economists, 5, 77–105.

Graff Zivin, J. and M. Neidell (2014): "Temperature and the allocation of time: Implications for climate change," Journal of Labor Economics, 32, 1–26.

Hanna, R. and P. Oliva (2015): "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City," Journal of Public Economics, 122, 68–79.

Heal, G. and J. Park (2016): "Reflections-temperature stress and the direct impact of climate change: A review of an emerging literature," Review of Environmental Economics and Policy, 10, 347–362.

Kahn, M. E. and P. Li (2019): "The effect of pollution and heat on high skill public sector worker productivity in China," NBER Working Paper.

Kumar, A. (2018): "Editorial: Neuroinflammation and Cognition," Frontiers in Aging Neuroscience, 10.

Landrigan, P. J., R. Fuller, N. J. R. Acosta, O. Adeyi, R. Arnold, N. N. Basu, A. B. Baldé, R. Bertollini, V. Fuster, M. Greenstone, A. Haines, D. Hanrahan, D. Hunter, M. Khare, A. Krupnick, B. Lanphear, B. Lohani, K. Martin, K. V. Mathiasen, M. A. Mcteer, C. J. L. Murray, J. D. Ndahimananjara, F. Perera, J. Potočnik, A. S. Preker, J. Ramesh, J. Rockström, C. Salinas, L. D. Samson, K. Sandilya, P. D. Sly, K. R. Smith, and A. Steiner (2018): "The Lancet Commission on pollution and health," Lancet, 391, 462–512.

Levitt, S. D., J. A. List, S. E. Sadoff, S. The, A. Economic, N. April, B. S. D. Levitt, J. A. List, and S. E. Sadoff (2011): "Checkmate : Exploring Backward Induction among Chess Players," American Economic Review, 101, 975–990.

Lichter, A., N. Pestel, and E. Sommer (2017): "Productivity effects of air pollution: Evidence from professional soccer," Labour Economics, 48, 54–66.

Liu, K., G. Sun, B. Li, Q. Jiang, X. Yang, M. Li, L. Li, S. Qian, L. Zhao, Z. Zhou, K. M. von Deneen, and Y. Liu (2013): "The impact of passive hyperthermia on human attention networks: An fMRI study," Behavioural Brain Research, 243, 220–230.

Meyer, S. and M. Pagel (2017): "Fresh Air Eases Work, The Effect of Air Quality on Individual Investor Activity," NBER Working Paper Series.

Palacios-Huerta, I. and O. Volij (2009): "Field centipedes," American Economic Review, 99, 1619–1635.

Park, J. (2018): "Hot Temperature, Human Capital and Adaptation to Climate Change," .

Pope, C. A., R. T. Burnett, G. D. Thurston, M. J. Thun, E. E. Calle, D. Krewski, and J. J. Godleski (2004): "Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution: Epidemiological Evidence of General Pathophysiological Pathways of Disease," Circulation, 109, 71–77.

Power, M. C., S. D. Adar, J. D. Yanosky, and J. Weuve (2016): "Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research," NeuroToxicology, 56, 235–253.

Roth (2018): "The Contemporaneous Effect of Indoor Air Pollution on Cognitive Performance: Evidence from the UK," Mimeo, 1–32.

Rückerl, R., A. Schneider, S. Breitner, J. Cyrys, and A. Peters (2011): "Health effects of particulate air pollution: A review of epidemiological evidence," Inhalation Toxicology, 23, 555–592.

Rustichini, A. (2015): "The role of intelligence in economic decision making," Current Opinion in Behavioral Sciences, 5, 32–36.

Satish, U., M. J. Mendell, K. Shekhar, T. Hotchi, D. Sullivan, S. Streufert, and W. J. Fisk (2012): "Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance." Environmental health perspectives, 120, 1671–7.

Schlenker, W. and W. R. Walker (2016): "Airports, Air pollution, and Contemporaneous Health," Review of Economic Studies, 83, 768–809.

Seppänen, O., W. Fisk, and Q. Lei (2006): "Effect of Temperature on Task Performance in Office Environment," Lawrence Berkeley National Laboratory, 11.

Seppänen, O. A. and W. J. Fisk (2004): "Summary of human responses to ventilation," Indoor Air, 14, 102–118.

Stafford, T. M. (2015): "Indoor air quality and academic performance," Journal of Environmental Economics and Management, 70, 34–50.

Stankovic, A., D. Alexander, C. M. Oman, and J. Schneiderman (2016): "A Review of Cognitive and Behavioral Effects of Increased Carbon Dioxide Exposure in Humans," NASA/TM-2016-219277, 1–24.

Taylor, L., S. L. Watkins, H. Marshall, B. J. Dascombe, and J. Foster (2015): "The Impact of Different Environmental Conditions on Cognitive Function: A Focused Review," Front Physiol, 6, 372.

Underwood, E. (2017): "The polluted brain," Science, 355, 342–345.

Xiaofang, Y., W. Rodney, Y. Weiwei, V. Pavla, P. Xiaochuan, and T. Shilu (2012): "Ambient temperature and morbidity: A review of epidemiological evidence," Environmental Health Perspectives, 120, 19–28.

Zhang, X., P. Wargocki, and Z. Lian (2015): "Effects of Exposure to Carbon Dioxide and Human Bioeffluents on Cognitive Performance," Procedia Engineering, 121, 138–142.

Zhang, Xin, Xi Chen and X. Zhang (2018): "The impact of exposure to air pollution on cognitive performance," Proceedings of the National Academy of Sciences, 115, 9193–9197.

Zivin, G. and M. Neidell (2012): "The impact of pollution on worker productivity," American Economic Review, 102, 3652–3673.

# Tables and Figures

Table 1: Descriptive statistics

| | N<br>(1) | mean<br>(2) | sd<br>(3) | min<br>(4) | max<br>(5) |
|---|---|---|---|---|---|
| *A. Player characteristics* | | | | | |
| *ELO* rating score | 102 | 1,681 | 329.1 | 950.3 | 2,289 |
| Number of official matches played | 101 | 80.83 | 64.12 | 2 | 273 |
| Age (in years) | 102 | 53.71 | 16.63 | 18 | 89 |
| Female | 102 | 0.0386 | 0.192 | 0 | 1 |
| *B. Game-specific characteristics* | | | | | |
| Total number of moves | 418 | 38.94 | 14.70 | 15 | 98 |
| Total duration (in minutes) | 413 | 171.50 | 54.62 | 43 | 310 |
| Draw game | 418 | 0.20 | | | |
| Player-opponent difference in | | | | | |
| *ELO* rating score | 418 | 3.51 | 357.60 | -1,265 | 814 |
| Experience (in #games) | 398 | 67.24 | 53.36 | 0 | 271 |
| Age (in years) | 418 | 18.13 | 14.24 | 0 | 66 |
| *C. Move-specific characteristics* | | | | | |
| Meaningful error | 20,408 | 0.08 | | | |
| Error if $> 0$ | 8,600 | 1.43 | 4.61 | 0.01 | 59.22 |
| *D. Environmental measures (round level)*$^{a)}$ | | | | | |
| CO2 (in ppm) | 14 | 1,549 | 326.60 | 1,179 | 2,393 |
| Temperature (in ℃ ) | 14 | 25.17 | 2.12 | 22.10 | 28.75 |
| PM2.5 (in $\mu g/m3$) | 14 | 22.38 | 9.15 | 14.03 | 51.05 |

*a)* Environmental measures are mean values of the prevailing conditions as measured during second hour of the tournament round.

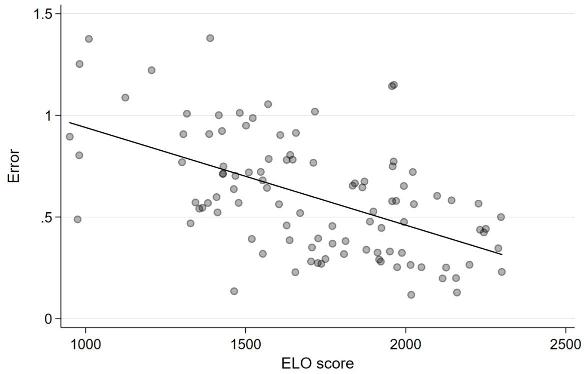
Table 2: Impact of indoor environmental quality on performance of chess players

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Panel A: Meaningful error* | | | | | |
| CO2 (in 100 ppm) | 0.000 | 0.003* | 0.003** | 0.002 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.003) | (0.003) |
| Temperature | 0.003 | -0.006** | -0.008** | -0.006 | -0.005 |
| | (0.003) | (0.003) | (0.003) | (0.005) | (0.005) |
| PM2.5 (in 10 $\mu g/m3$) | 0.004 | 0.011** | 0.011** | 0.013*** | 0.015*** |
| | (0.005) | (0.004) | (0.004) | (0.005) | (0.005) |
| Observations | 20,408 | 20,408 | 20,408 | 20,408 | 20,408 |
| Adj. R-squared | 0.009 | 0.018 | 0.018 | 0.019 | 0.037 |
| *Panel B: Ln(error)* | | | | | |
| CO2 (in 100 ppm) | -0.011 | 0.006 | 0.008 | 0.016 | -0.001 |
| | (0.013) | (0.012) | (0.012) | (0.021) | (0.022) |
| Temperature | 0.045* | -0.015 | -0.029 | -0.043 | -0.022 |
| | (0.025) | (0.024) | (0.026) | (0.041) | (0.043) |
| PM2.5 (in 10 $\mu g/m3$) | 0.011 | 0.051 | 0.054 | 0.073* | 0.094** |
| | (0.044) | (0.036) | (0.036) | (0.042) | (0.045) |
| Observations | 8,600 | 8,600 | 8,600 | 8,600 | 8,589 |
| Adj. R-squared | 0.024 | 0.050 | 0.050 | 0.051 | 0.113 |
| Player FE | NO | YES | YES | YES | YES |
| Tournament FE | NO | NO | YES | YES | YES |
| Round FE | NO | NO | NO | YES | YES |
| Move FE | NO | NO | NO | NO | YES |

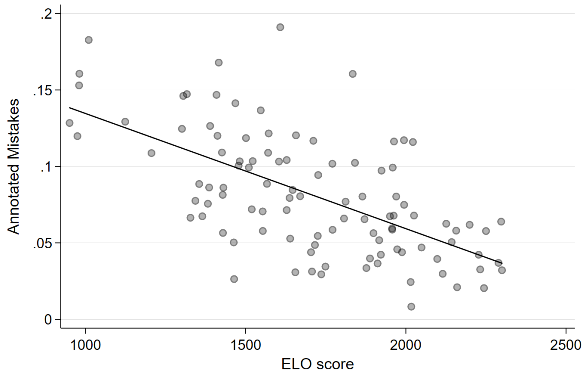
*Note:* */**/*** indicate statistical significance at the 10%/5%/1% levels. Standard errors are in parentheses and clustered at the game level. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes the value of 1 if the move is marked as a meaningful error by the chess engine and zero otherwise. For each panel, each column displays the results of a separate regression with the combination of fixed effects specified at the bottom of the table. All regressions presented in the table include all the environmental parameters and the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely, the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).

Figure 1: Player skills and average move performance

Panel A. Average error and player *ELO* rating score




Panel B. Average number of meaningful errors and player *ELO* rating score)




*Note*: Each dot in the figures represents a player, the figures display the average error of a player (Panel A) or the average number of annotated errors (Panel B) in the vertical axis, and the average *ELO* rating score of the player over the two tournaments in the sample. The error measure is defined in equation 1. The annotated errors are defined as the sum of moves labeled as mistakes ('?') and blunders ('??'). The Pearson correlation between the error measure and the *ELO* rating score is -0.54 (p-value=0.00). The Pearson correlation between the average number of annotated errors and the *ELO* rating score is -0.62 (p-value=0.00). The correlation between the player average of two move-performance measures is 0.72 (p-value=0.00).

Figure 2: Distribution of total number of moves per game

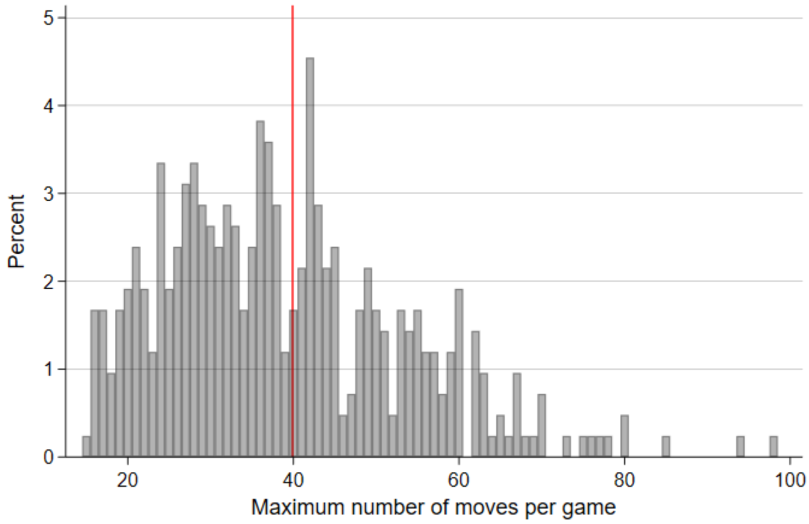
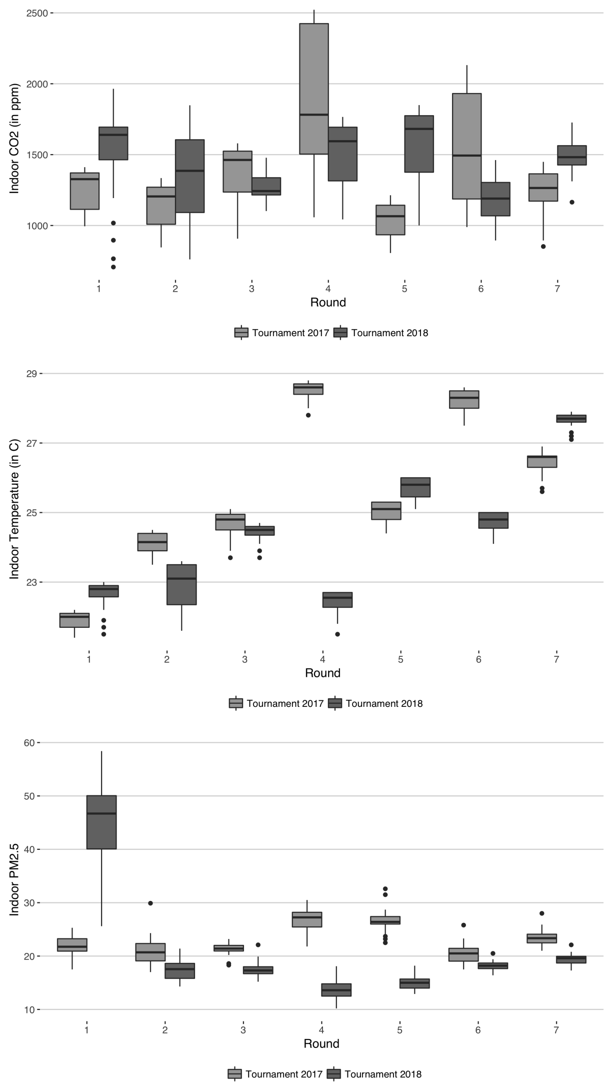
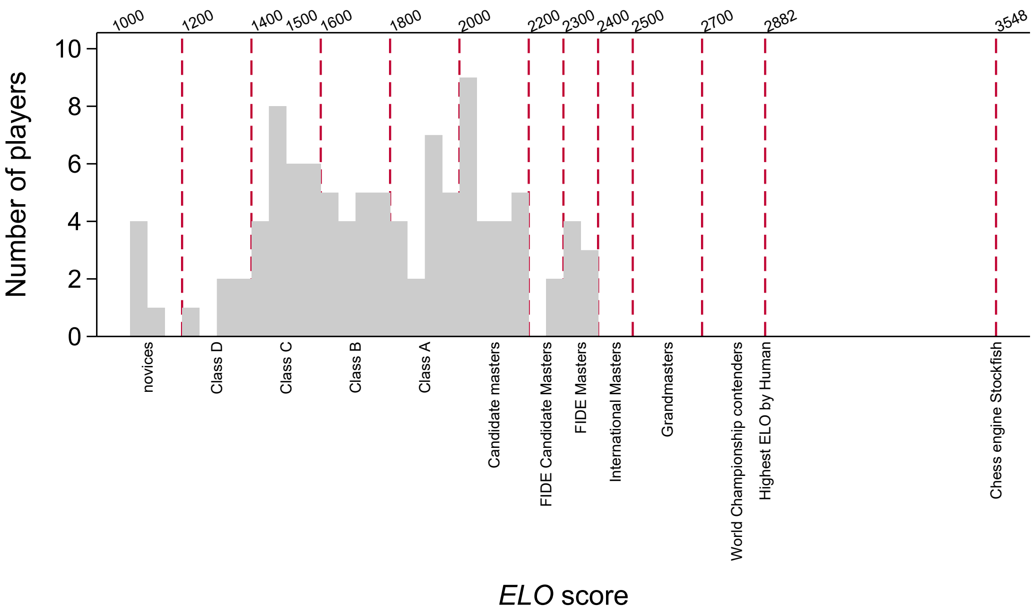



Figure 3: Indoor environmental conditions as measured on the days at the tournaments



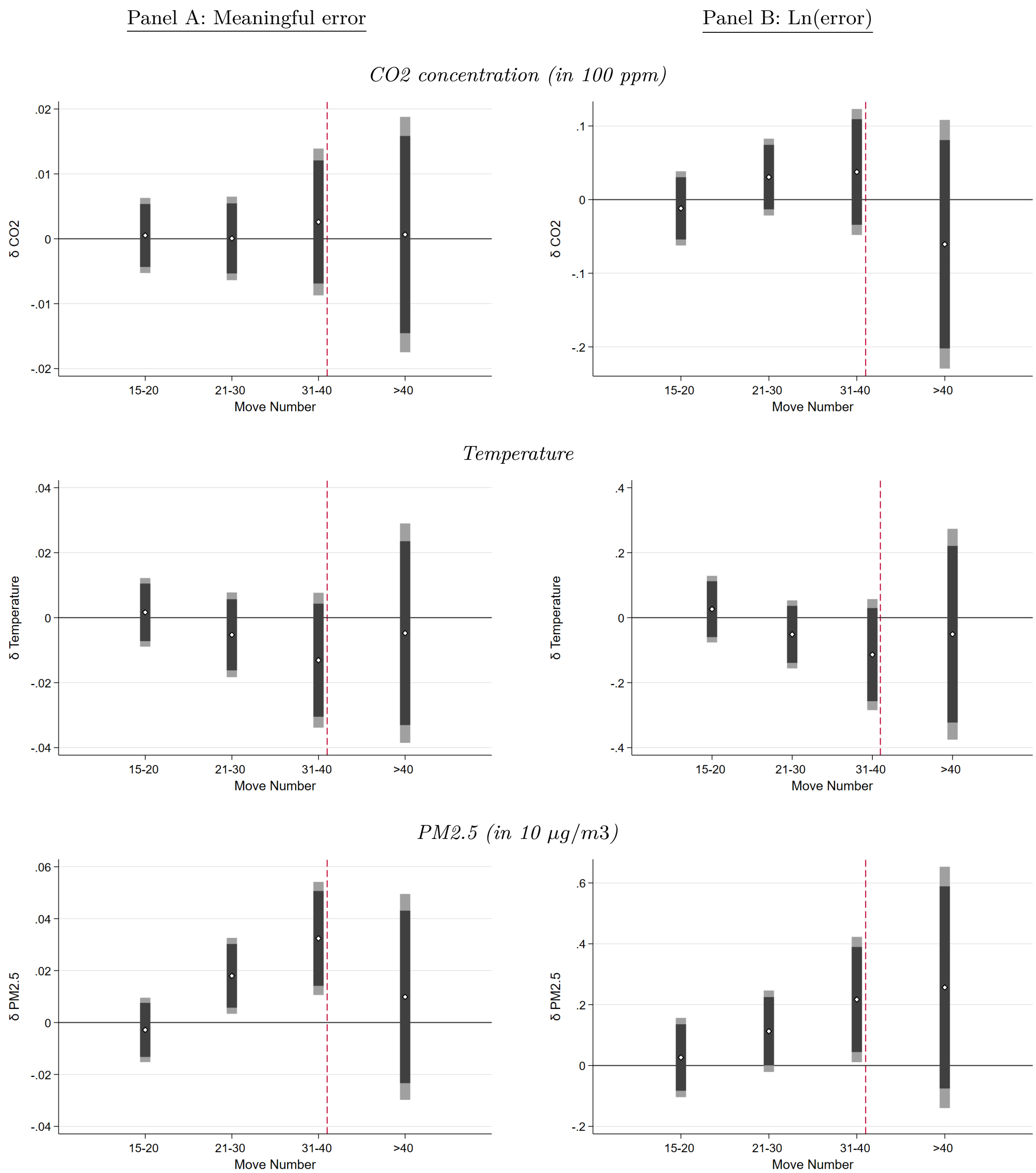

*Note:* The figures show the distribution of CO2 concentration, temperature, and fine particulate matters (PM2.5) are measured during the days (rounds) at the chess tournaments.

Figure 4: Distribution of players' *Elo* rating score




*Note:* The players' Elo score is calculated by adding 100 to the players' DWZ score in order to make the scores comparable to the FIDE system.

Figure 5: Impact of indoor environmental quality on performance of chess players by move level

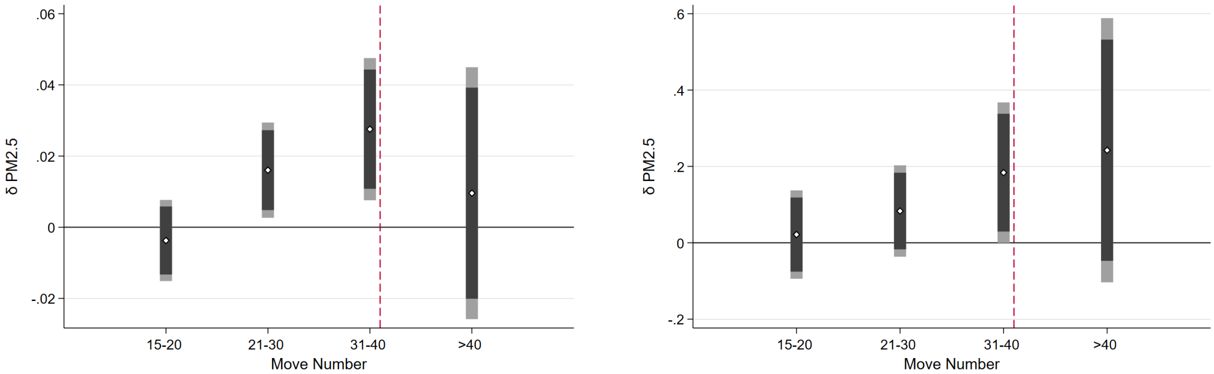



*Note:* The figure shows the estimated coefficient of joint regressions including all the environmental measures. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The vertical, dashed red line indicates the occurrence of the time restriction during the game. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes the value of 1 if the move is marked as a meaningful error by the chess engine, and zero otherwise. Dots represent point estimates. Black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely., the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).

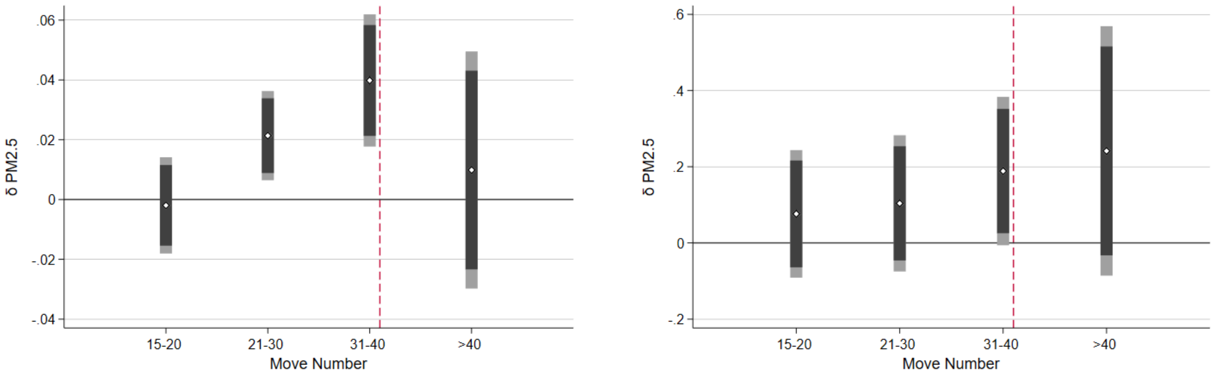
Figure 6: Robustness of the effect on PM2.5

Meaningful error | Ln(error)

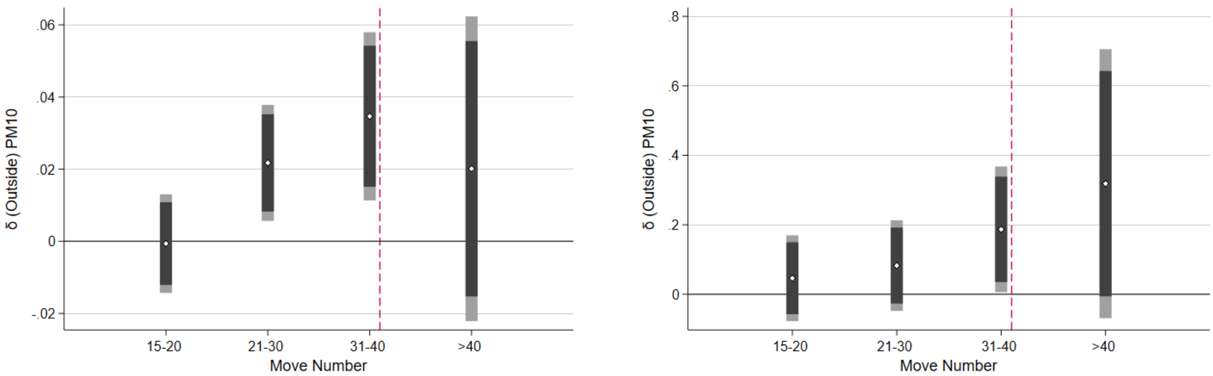
*Panel A: Daily maximum value of air quality as treatment*




*Panel B: Only games with $\geq 40$ moves per player*




*Panel C: Outdoor PM10 (in 10 $\mu g/m3$)*

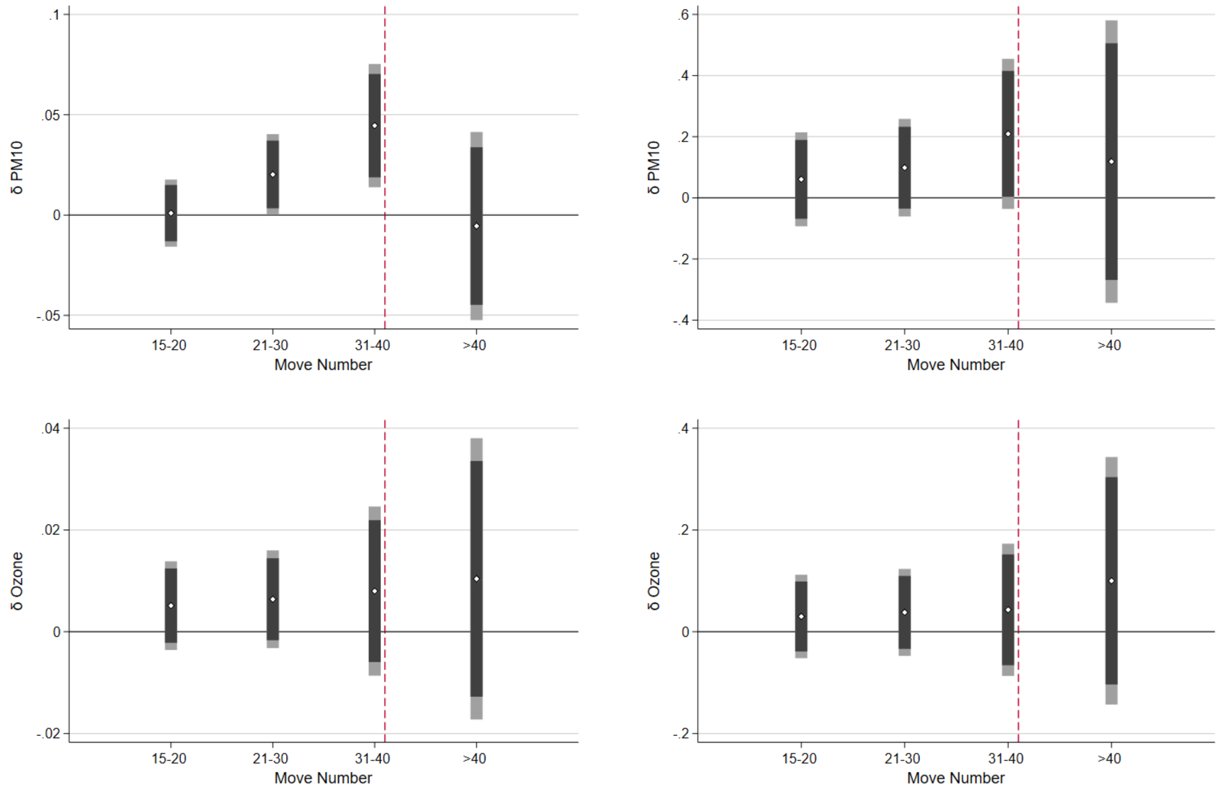



*Note:* The figure shows the results of the sensitivity analysis testing the robustness of the effect on PM2.5. We show the estimated coefficient of joint regressions including all the environmental measures. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The vertical, dashed red line indicates the occurrence of the time restriction during the game. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes the value of 1 if the move is marked as a meaningful error by the chess engine, and zero otherwise. Dots represent point estimates. Black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely, the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).

Figure 7: Robustness Check Inclusion of Ozone

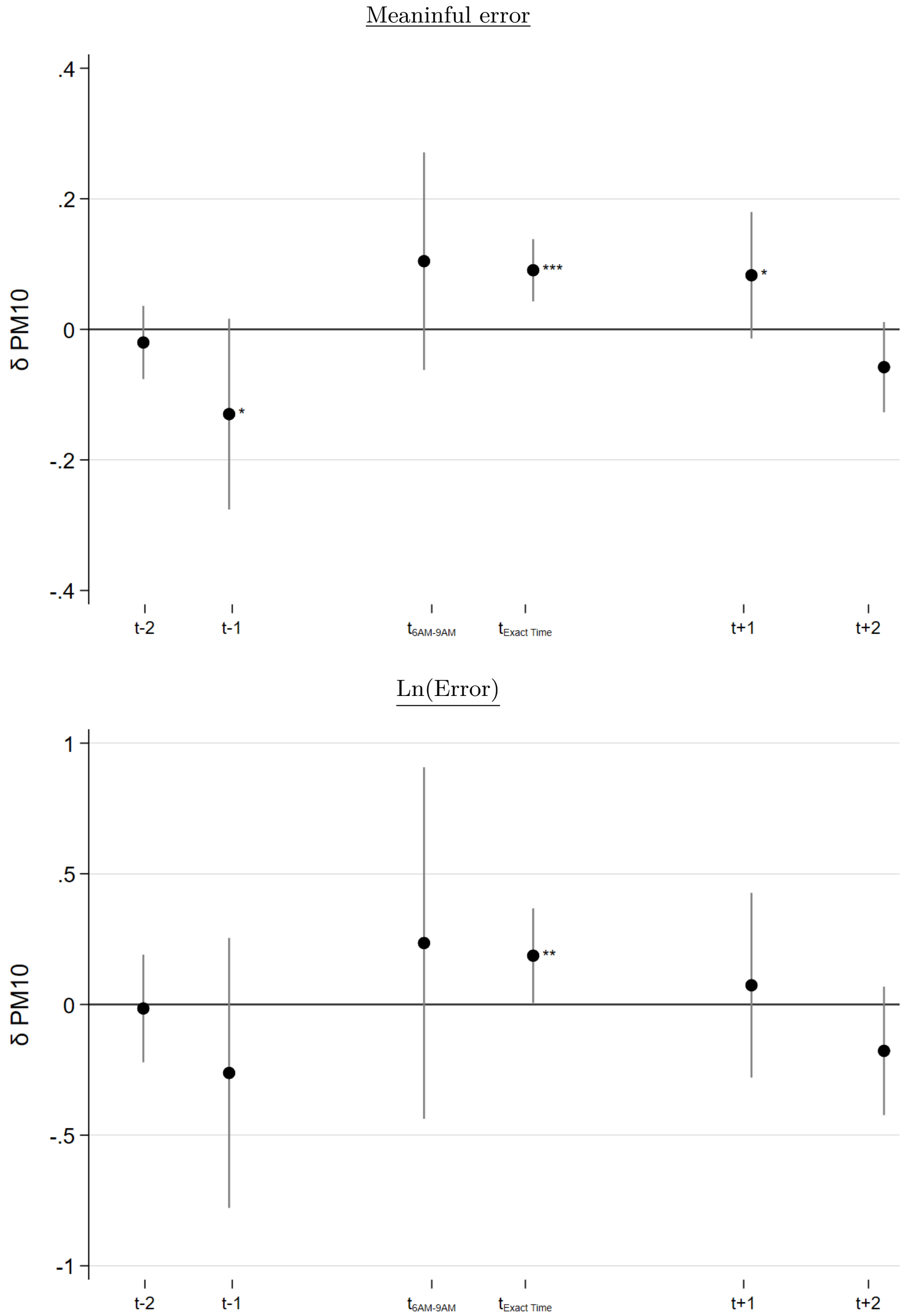
Meaningful error

Ln(error)




*Note:* We show the estimated coefficient of joint regressions including all the environmental measures. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The vertical, dashed red line indicates the occurrence of the time restriction during the game. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes the value of 1 if the move is marked as a meaningful error by the chess engine, and zero otherwise. Dots represent point estimates. Black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely, the pawn metric of the previous move by the opponent ($C_{jtrm-1}^{opponent}$).

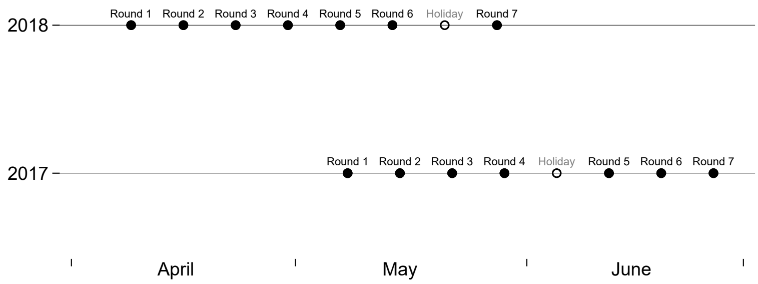
Figure 8: Falsification Test




*Note:* */**/*** indicate statistical significance at the 10%/5%/1% levels. We show the estimated coefficient of separate regressions. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The binary outcome variable ”meaningful error” takes the value of 1 if the move is marked as a meaningful error by the chess engine and zero otherwise. Dots represent point estimates. Gray bars show the 95% confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely., the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).

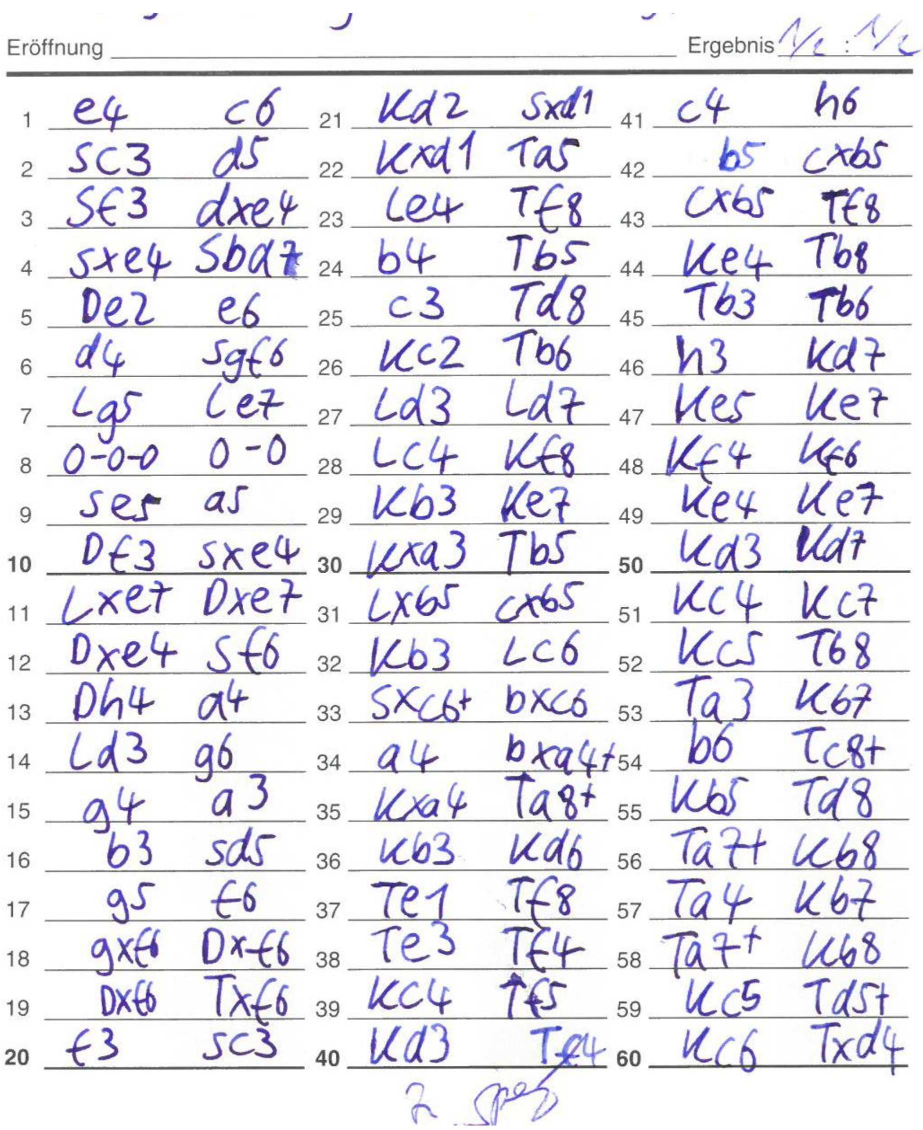
# A Appendix

Figure A.1: Timing and setting of the chess tournaments




*Note:* This diagram illustrates the timing and setting of the observed tournaments. Each tournament consists of seven rounds, played every Monday, 6:00pm (local time).

Figure A.2: Example of players' hand-written game notation



Eröffnung Ergebnis 1/2 : 1/2

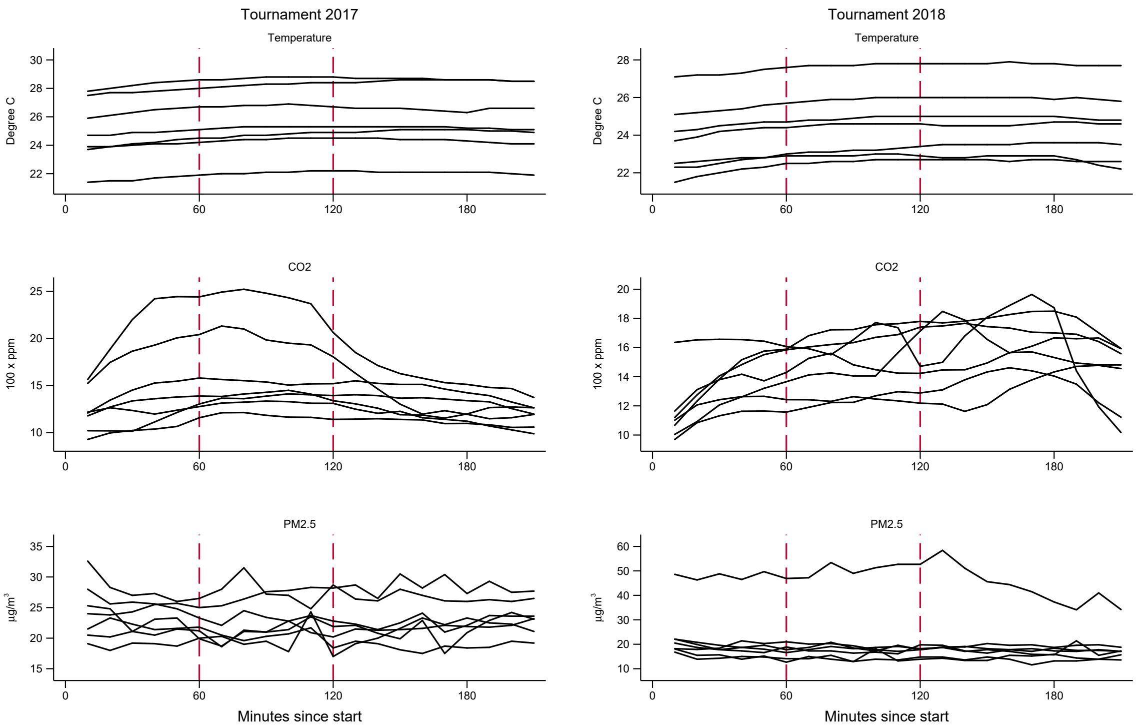
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | e4 | c6 | 21 | Kd2 | Sxd1 | 41 | c4 | h6 |
| 2 | Sc3 | d5 | 22 | Kxd1 | Ta5 | 42 | b5 | cxb5 |
| 3 | Sf3 | dxe4 | 23 | Le4 | Tf8 | 43 | cxb5 | Tf8 |
| 4 | Sxe4 | Sbd7 | 24 | b4 | Tb5 | 44 | Ke4 | Tb8 |
| 5 | De2 | e6 | 25 | c3 | Td8 | 45 | Tb3 | Tb6 |
| 6 | d4 | Sgf6 | 26 | Kc2 | Tb6 | 46 | h3 | Kd7 |
| 7 | Lg5 | Le7 | 27 | Ld3 | Ld7 | 47 | Ke5 | Ke7 |
| 8 | 0-0-0 | 0-0 | 28 | Lc4 | Kf8 | 48 | Kf4 | Kf6 |
| 9 | Se5 | a5 | 29 | Kb3 | Ke7 | 49 | Ke4 | Ke7 |
| 10 | Df3 | Sxe4 | 30 | Kxa3 | Tb5 | 50 | Kd3 | Kd7 |
| 11 | Lxe7 | Dxe7 | 31 | Lxb5 | cxb5 | 51 | Kc4 | Kc7 |
| 12 | Dxe4 | Sf6 | 32 | Kb3 | Lc6 | 52 | Kc5 | Tb8 |
| 13 | Dh4 | a4 | 33 | Sxc6+ | bxc6 | 53 | Ta3 | Kb7 |
| 14 | Ld3 | g6 | 34 | a4 | bxa4+ | 54 | b6 | Tc8+ |
| 15 | g4 | a3 | 35 | Kxa4 | Ta8+ | 55 | Kb5 | Td8 |
| 16 | b3 | Sd5 | 36 | Kb3 | Kd6 | 56 | Ta7+ | Kb8 |
| 17 | g5 | f6 | 37 | Te1 | Tf8 | 57 | Ta4 | Kb7 |
| 18 | gxf6 | Dxf6 | 38 | Te3 | Tf4 | 58 | Ta7+ | Kb8 |
| 19 | Dxf6 | Txf6 | 39 | Kc4 | Tf5 | 59 | Kc5 | Td5+ |
| 20 | f3 | Sc3 | 40 | Kd3 | Tf4 | 60 | Kc6 | Txd4 |

*Note:* This picture shows an example of the hand-written documentation as filled in during each game within the chess tournament. The documentation has been digitized by the tournament organizers.

Figure A.3: Example for sensor location

*Note:* This picture illustrates the placement of one sensor measuring the indoor environmental quality. In total, three sensors were placed across the room on separate tables.

Figure A.4: Distribution of indoor-environmental-quality measures during the tournament rounds

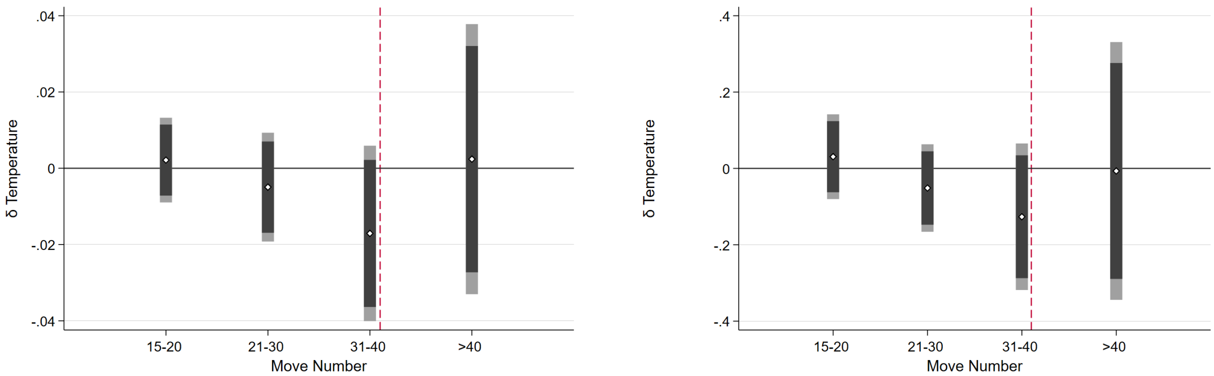



*Note:* The solid black lines indicate the distribution of the environmental measures during the seven rounds within a tournament. The calculation of the mean values of the environmental measures as used in the regression analysis are calculated based on observations during the second hour of the tournament rounds, as indicated by the dashed lines.

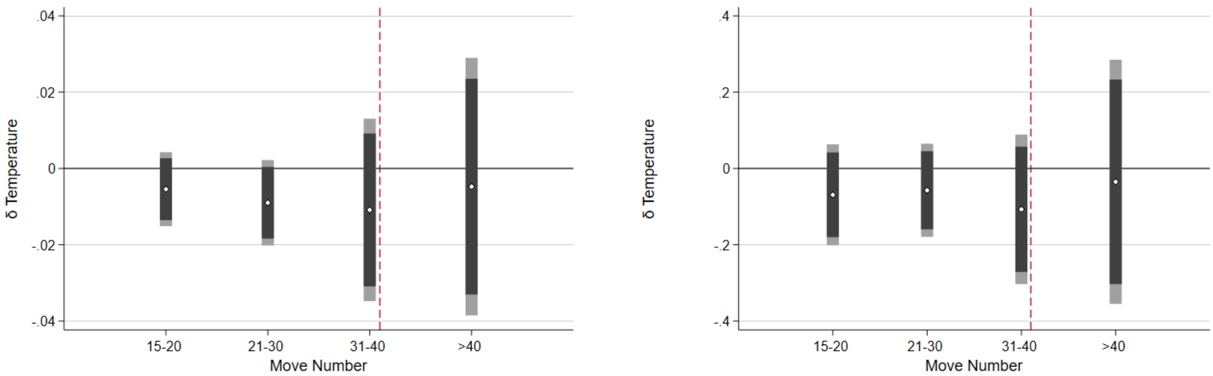
Figure A.5: Robustness of the effect on temperature

Meaningful error | Ln(error)

*Panel A: Daily maximum value of temperature and air quality as treatment*




*Panel B: Only games with $\geq 40$ moves per player*

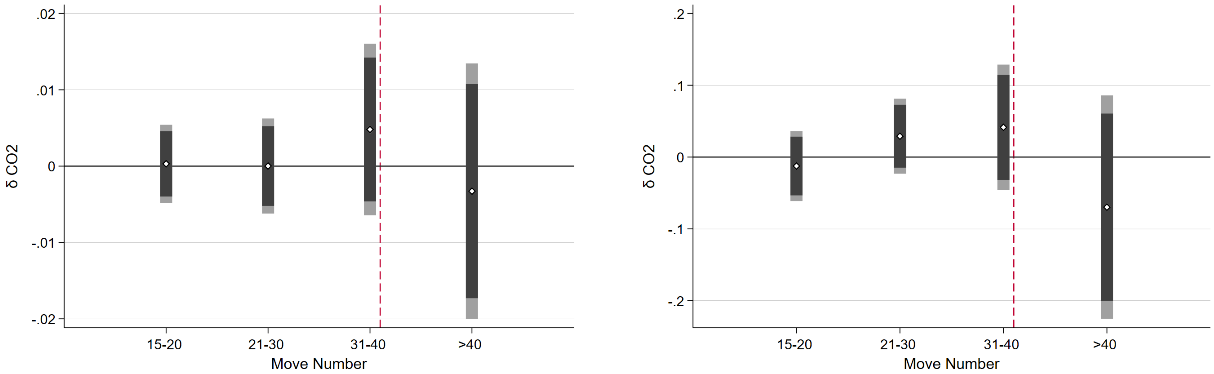



*Note:* The figure shows the results of the sensitivity analysis testing the robustness of the effect on PM2.5. We show the estimated coefficient of joint regressions including all the environmental measures. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The vertical, dashed red line indicates the occurrence of the time restriction during the game. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes the value of 1 if the move is marked as a meaningful error by the chess engine, and zero otherwise. Dots represent point estimates. Black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely, the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).

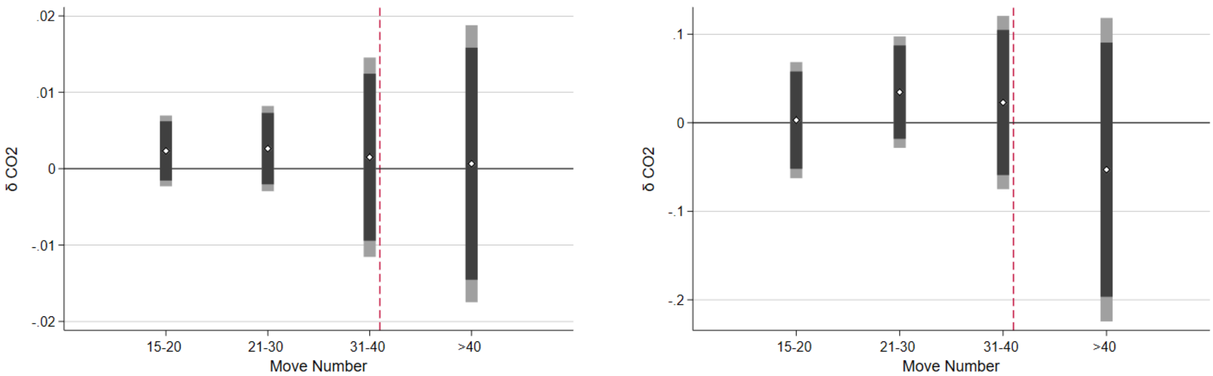
Figure A.6: Robustness of the effect on CO2

Meaningful error                                    Ln(error)

*Panel A: Daily maximum value of temperature and air quality as treatment*




*Panel B: Only games with $\geq 40$ moves per player*




*Note:* The figure shows the results of the sensitivity analysis testing the robustness of the effect on PM2.5. We show the estimated coefficient of joint regressions including all the environmental measures. We divided the total sample of moves into subsamples with respect to the number of moves within a game (horizontal axis). The vertical, dashed red line indicates the occurrence of the time restriction during the game. Each panel presents the regression on different outcomes. The binary outcome variable "meaningful error" takes of the value of 1 if the move is marked as a meaningful error by the chess engine, and zero otherwise. Dots represent point estimates. Black (gray) bars show the 90% (95%) confidence intervals based on standard errors clustered at the game level. All regressions include individual, year, round, and move fixed effects, as well as the full set of control variables: (i) difference in the *ELO* rating score between the player and the opponent (as well as its squared term), (ii) the number of points achieved during the tournament, and (iii) the actual status of the game before the move, namely., the pawn metric of the previous move by the opponent ($C^{opponent}_{jtrm-1}$).