Clouded Thoughts: Air Quality and Cognitive Performance*

Arthur Novaes de Amorim†

October 17, 2018

Abstract

This paper investigates the impact of air pollution on the quality of decisions in a bounded rationality setting. I exploit randomly occurring dust storms to quantify the effect of air pollution on cognitive performance of players of the strategy board game Go. I benchmark the quality of moves played by humans against move evaluations from a powerful Go artificial intelligence. My results show that particulate matter (PM) exposure increases blundering in game. This effect grows almost linearly with age; a 30 year old player experiences 15% more blunders with a PM shock of $75 \mu g/m^3$, corresponding to a typical dust storm. Age asymmetric effects translate to game outcomes favoring younger players during dust storm days.

NOTE: VERY PRELIMINARY VERSION. NOT TO BE CITED WITHOUT AUTHOR'S PERMISSION.

Keywords: Air quality, decision-making
JEL Codes: D87, Q53, Q54

*Acknowledgments
†University of Calgary.

1 Introduction

Recent deterioration in air quality due to economic growth is a concern in many developing economies. While news headlines often emphasize the health effects of air pollution (e.g. "WHO reveals 7 million die from pollution each year [...]", 2018), evidence suggests that air quality may also affect worker productivity. This strand of the literature gained momentum with Graff Zivin and Neidell (2012), who show that increased Ozone (O_3) and fine particulate matter ($PM_{2.5}$) exposure decreases the productivity of fruit pickers and packers. Arguably, most high value jobs require cognitive abilities not cultivated by physical labourers. My research investigates how air quality affects one such cognitive ability in a particular context, namely the decision-making of individuals engaging in strategic interactions.

Estimating a relationship between air quality and cognitive performance poses econometric challenges. First, as pointed out in Lavy et al. (2014), air pollution often correlates with cognitive performance through factors such as per-capita income since well-paid high-skill workers may sort into cleaner locations. Second, while an objective and reliable metric of decision-making is necessary for estimation, performance assessments in many cognitively-demanding tasks are subjective. My empirical strategy attempts to overcome these issues and estimates a *causal* effect of air pollution on the quality of decision-making for expert players of the game Go.

Go is a strategy board game, in which two players take turns placing coloured stones on the vacant intersections of a board, where broadly the objective is to surround more territory than the opponent. The game recently caught public attention when the first-ranked player in the world was defeated by an artificial intelligence (AI) developed by Google DeepMind. I analyze records of high-level Go games using Leela Zero[1] – an open-source AI modeled after DeepMind's Alpha Go Zero – as an "expert evaluator" which classify players' moves according to their quality. Leela Zero has a strong record of defeating professional Go human players with generous handicaps against the AI, and its strength is estimated to surpass the current world champion.[2]

I ask Leela Zero to evaluate moves from historical games played by professionals and highly skilled amateurs and use these evaluations to construct for each move an objective measure for the quality of decision-making of Go players. The game has a long history of competitive play in Korea and Japan where many tournaments are played every year,

[1]The term Zero indicates that the AI was trained on a neural network without human inputs, as opposed to the original Alpha Go which learned to play from records of expert players.

[2]See for example the series Leela Zero Vs. Haylee, where Leela Zero defeated the 4-dan professional Hajin "Haylee" Lee in 8 out of 8 games. The games are available at Haylee's YouTube channel *Haylee's World of Go/Baduk."*

providing decades of annotated game data. Tournaments are typically played indoor in hotels and game salons, in controlled environments which mitigates threats to identification like temperature and weather conditions. The indoor settings also imply I estimate the effects on cognitive performance of air pollution with high indoor penetration factor, such as particulate matter (PM).[3]

In this paper, I construct a dataset that exploits shocks in air pollution exposure of Go players in Korea and Japan by matching the day and location of games with daily regional records of events known as Asian dust storms. Asian dust storms are well documented natural phenomena responsible for transporting PM and other pollutants from Mongolia and Northern China to neighbouring eastern countries through jet streams. These dust events occur sporadically, and are a growing environmental concern due to rising pollution levels in mainland China (Mosteller, 2016). Moreover, the random nature of the Asian dust means it can be exploited as a source of exogenous variation in air pollution exposure to estimate a causal effect on cognitive performance.

Equipped with the data described above, I answer the following questions: (1) is the quality of decision-making of Go players affected by substantial changes in air pollution? If so, (2) is such an effect heterogeneous across observable characteristics such as player age and skill level? The answers to these questions improve our understanding on the benefit side of clean air policies. Below I present my methods and results.

First, I establish that the metrics of cognitive performance constructed from Leela Zero's move evaluations are predictive of a player's strength and of the game outcome. Stronger players match the AI-recommended move more often then weaker players. Conversely, weaker players blunder more frequently. Moreover, the probability of winning the game increases sharply for one additional move matching the AI (or one fewer blunder) relative to the opponent, even after controlling for each player's strength and demographic characteristics. Second, I document the relationship between Asian dust and pollution levels in South Korean and Japanese cities. During Asian dust days, measured levels of different particulate matters (suspended SPM, coarse PM_{10}, and fine $PM_{2.5}$) increase between 45 to 75% in South Korean and Japanese cities. To put in perspective, the metropolis Seoul in South Korea registered 25 dust storms in 2001, and average PM_{10} levels for that year are $75\mu g/m^3$ higher during dust days.[4] Other pollutants, namely O_3, SO_2 and CO, see only a modest increase between 3 and 8% during dust days, which suggests that particulate matter pollution is a key driver

[3]Papers such as Ozkaynak et al. (1996) suggest PM penetrates buildings through physical openings as well as ventilation systems. Also, evidence from the health literature links PM with central nervous system disorders such as migraine, headache, and stroke (Loane et al., 2013).

[4]One study reveals an increase of 22% in lung cancer for every $10\mu g/m^3$ increase in PM_{10}.

of the main results below.[5]

Finally, using a fixed effects model I estimate the effect of Asian dust on the cognitive performance outcomes. My main finding is that Asian dust has a large and significant effect on a player's propensity to blunder. My estimates show that players overall make 6.5% more blunders when a game is played during a dust day, which amounts to roughly two additional blunders during a standard game lasting 200 moves. For a rough comparison: recent work by Archsmith et al. (2018) find baseball umpires make one additional incorrect "ball/strike" call per each 250 decisions when exposed to an additional $10\mu g/m^3$ in 12-hour $PM_{2.5}$.

I uncover some heterogeneity after reproducing these estimates for the subpopulations below and above 30 years of age; the Asian dust effect dissipates for younger players and becomes more pronounced for older players, implying the older players make 14.7% more blunders during the induced air pollution shock. These heterogeneous effects are consistent with a strand of the health literature suggesting older individuals are more susceptible to adverse effects of air pollution. I also find heterogeneous effects by player strength: lower-ranked professionals and amateurs – which I argue are more likely to resemble other decision-makers in the population – make on average four additional blunders per game.

In contrast, I do not find statistically significant Asian dust effects on the players' propensity to make strong moves, neither in the full sample nor in the age and rank groups. The point estimates are fairly precise and consistently close to zero across all regression specifications except for the subpopulation of older players, which has coefficients with a larger magnitude (but still insignificant). An interpretation of these results is that air pollution leads to an increase in the quantity of poor decisions but not a decrease in the quality of good decisions of Go players.

This paper complements the literature concerning effects of air quality on cognitive performance of decision makers. Focusing on Go allows me to extend this body of research by documenting a significant impact of air pollution on a *purely* cognitive task which demand a high degree of inductive reasoning and is performed in a laboratorial environment. While the peculiarity of Go may seem to limit the extent to which this contribution generalizes, the narrow set of cognitive functioning used by the game players lead to clean identification, i.e. this contribution speaks of tasks requiring a high level of inductive reasoning.[6]

[5] A variety of toxic materials are found in Asian dust and I cannot rule out the possibility that a different pollutant is behind the research findings. However, the Japanese data contains a few additional pollutants which I find to be unaffected by dust storms.

[6] Go belongs to a class of non-trivial (i.e. "hard to solve") combinatorial games which also includes chess and checkers. Experts in these games are known to have a high degree of inductive reasoning and often gather research interest. See for example Levitt et al. (2011) and Palacios-Huerta and Volij (2009) for experimental tests of chess players ability to backward induct in games such as Centipede and Race to 100. See Biswas and Regan (2015) for empirical work relating to k-level thinking and satisficing among chess players.

Asian dust has been featured on articles by the New York Times (French, 2002) and Reuters (Herskovitz, 2008) as an environmental problem choking economic growth in South Korea and Japan. From a policy perspective, my results also contribute to the discussion of Chinese pollution spillovers to neighbouring countries by proposing a new channel in which Asian dust may affect worker productivity.

Lastly, the data construction complements a growing literature on measuring worker performance with off-the-shelf machine learning (ML) algorithms (see Chalfin et al. (2016) for a recent application of ML on predicting labour productivity). Until recently, relating cognitive performance to move choices in a board game seemed to be a daunting task due to computational limitations and algorithmic complexity. To my knowledge, two current research teams have recently tackled a similar task: Biswas and Regan (2015) relates chess moves to k-level thinking using the depth of search feature in a chess AI and Backus et al. (2016) use ELO rating estimates from a chess engine to measure game quality of play.

The remainder of the paper is organized as follows. Section I reviews the literature on health effects of Asian dust and the relationship of air quality and economic growth. Section II provides a background on the history of Go and the recent advances in Go-playing AI. Section III describes the air quality data and database of game records. Section IV outlines the empirical strategy used to estimate the effect of Asian dust on strategic thinking. In section V, I present and discuss the estimation results. Section VI concludes the paper.

2 Literature Review

2.1 Air Quality, Health, and Labour Productivity

There exists a vast literature that study the relationship between the environment and the population well-being.[7] This line of research has produced compelling evidence that air pollution adversely affects human health and subsequently impacts labour market outcomes on an extensive margin. An early example of such evidence is Hausman et al. (1984), who finds that a standard deviation increase in total suspended particulates is associated with an approximately ten percent increase in work days lost. A Chay and Greenstone (2003) provide a methodological contribution by exploiting geographic variation in air pollution in the US due to county-level income shocks induced by a recession. They find that a 1% reduction in total suspended particles results in a 0.35% decline in infant mortality rate at the country level.

Research capturing intensive margin effects of air pollution on worker productivity have

[7] Graff Zivin and Neidell (2013) and Currie et al. (2014) provide excellent reviews of this literature.

gathered academic interest in recent years. A key contribution to this strand of the literature, Graff Zivin and Neidell (2012) find strong evidence that short-term exposure to $PM_{2.5}$ and O_3 diminishes the productivity of fruit pickers working on Californian farms. The authors provide a back-of-the-envelope calculation suggesting that a 10ppb reduction in the ozone standard recommended by the EPA at the time would translate into annual savings of approximately \$700 million in labour expenditure. This contribution, while economically relevant (agriculture is particularly important in the developing world), has little implications for workers engaged in tasks that are mostly, or purely, cognitive.

Two recent papers investigating the impact of pollution exposure on different mental faculties are closely related to my research. First, Lavy et al. (2014) demonstrates that short-term exposure to air pollution adversely affects cognitive performance measured by student test scores. The authors exploit daily variation in the $PM_{2.5}$ exposure of a student during the days of writing the Bagrut, a series of high school exit exams used for university admissions and find transitory $PM_{2.5}$ exposure to significantly reduce test scores. Lowered test scores due to pollution exposure is found to decrease long-term educational outcomes and earnings. They also speculate air pollution may be more damaging for students with health conditions such as asthma after identifying more pronounced effects on a demographic group with higher incidence of respiratory illnesses. This heterogeneous effect could have long-lasting effects since it may lead "healthy" students with lower human capital to be matched with better schooling outcomes than more qualified "unhealthy" peers. In sum, their work explore the effects of air quality on a fairly broad measure of cognitive ability (university test scores) for an economically relevant population of students. My research complements Lavy et al. (2014) by disentangling adverse effects of air quality on a specific cognitive functioning, namely the decision-making of individuals known to perform inductive reasoning. In addition, the population of Go players in the data has a wide age distribution conducive for identifying differential effects of air pollution for distinct demographic groups.

Second, (Archsmith et al., 2018) provides evidence that air pollution negatively affects quality of "snap decisions" of umpires in a sports context. Their research shows that a $10\mu g/m^3$ short-term increase in 12-hour $PM_{2.5}$ exposure causes baseball umpires to make 2.6% more incorrect "ball/strike" calls. While the work of an umpire is not only cognitive but also physical, the task they perform is certainly quality-focused and requires a high degree of concentration. Mistakes in arbitrating baseball games may not be equivalent to blunders in strategy board games, but they nonetheless provide useful estimates against which I can benchmark my results.

2.2 Asian Dust Storms in South Korea and Japan

Asian dust is a natural phenomenon, which typically occur between September and May of each year, whereby dust particles from desert areas in Northern China and Mongolia are transported for long distances via jet streams. Historical records of a yellow dust traveling specifically from the Gobi desert to the Korean peninsula can be traced back to the year 174 A.D. (Chun et al., 2008). In recent years, however, the Asian dust carries a growing amount of major pollutants including PM_{10}, nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO) which are likely originated from China Lee et al. (2007). In addition to carrying local pollution from China, Mori et al. (2003) finds that, while it traverses from China to its neighbouring countries, this dust collects nitrate and sulphate ions thus generating other chemical compounds that cause negative health effects.

A body of literature on public health have studied the effects of Asian dust on mortality rates. Kwon et al. (2002) examine the effects of 28 dust events occurring in Seoul between 1995-1998 to find that death rates during Asian dust increased 4.1% for cardiovascular and respiratory causes, and the elderly subpopulation was the most affected by these adverse health effects. Lee et al. (2013) additionally document that the air pollution shock induced by Asian dust increased between 1995 and 2009, partly due to a reduction in local pollution level in major South Korean cities. Their research design exploits the implementation by the Korean government of public dust warnings and find suggestive evidence of a behavioural response where mortality effects decrease due to dust advisory.

Jia and Ku (2015) investigate whether pollution from China spills over to neighbouring countries through Asian dust storms. To do so, the authors propose a model that exploits spatial and over-time variation in dust incidence within South Korea with temporal variations in air quality in China. Their finding, after controlling for the direct effects of the Asian dust, links increases in pollution levels in China to higher mortality rates due to respiratory and cardiovascular diseases in South Korea, with the most prominent effects again observed on the elderly subpopulation.

While there is strong evidence supporting the health-related effects of the Asian dust, no attention has been given to the possibility that this phenomenon may affect economic outcomes through short-term deterioration of cognitive functioning. PM pollution can penetrate into lungs and, if the particulate is sufficiently fine, enter the bloodstream. These particulates originate from various sources such as automobile emissions and industrial activity. It is suggested that certain components of $PM_{2.5}$ may affect an individual's central nervous system and ultimately the brain. Loane et al. (2013) reviews this line of research and documents a positive association between PM and migraine, headache, stroke, Alzheimer's disease, and Parkinson's disease. Ghio et al. (2000) finds that even short-term exposure to PM may lead to mild conditions such as irritation in throat and lungs, with symptoms occasionally arising hours after exposure takes place. Genc et al. (2012) report by surveying experimental studies that long-term exposure to air pollution has a negative impact on the neural development in children after adjusting for socio-economic status, smoking, and blood lead levels. In light of these results, documenting the relationship of Asian dust pollution and cognitive performance seems warranted.

3 Go

3.1 A Primer on the Game

Go is a popular board game in East Asia, played on a 19×19 grid between two players, denoted Black and White, who alternately place coloured stones on the vacant board intersections. Players score points both by surrounding board territory with their stones and by capturing the opponent's stones. Capturing takes place when a stone (or a group of stones) is surrounded by opposing colour stones on all orthogonally-adjacent intersections. Players are allowed to pass a turn, and the game concludes when both players sequentially pass. The winner is determined by counting one point for each territory and captured stone of each player's and adding compensation points given to White for Black's first mover advantage [8].

While the game rules are simple, playing Go is a mentally taxing activity. The average number of move choices available per turn is over 200, significantly more than the chess average of 35 choices (Keene and Levy, 1992). Playing the game at a high level requires multiple cognitive processes such as attention, memory, and reasoning. Given the large average number of move choices, memory and pattern recognition with respect to spatial positioning may be more important than in other strategy games. As suggested by Gobet et al. (2004), mental tasks required in board games such as pattern recognition and memory for domain material "have been shown to generalize to most, if not all, domains of expertise."

3.2 Game records

The analysis uses data from historical Go game records published in printed and online media and compiled by the Games of Go on Disk (GoGoD) database. Each game is stored in a text file using a protocol called Smart Game Format (SGF) which records the entire sequence of moves in a game tree-based representation. The format allows the input of multiple

[8]Different ruling systems concurrently exist in various countries. The game records in this analysis mostly follow the 1989 Japanese revised rules, available in details at `http://www.cs.cmu.edu/~wjh/go/rules/Japanese.html`.

Table 1: Variables Extracted from Game Records, Player Biographies, and Ratings website

Property name	**Description**
Name	name of player
Stone Colour	player controls black or white stones
Rank	rank of player at game day
Age	age of player at game day
Sex	birth sex of player
Elo Rating	calculated rating of player at game day
# of Moves	number of moves played in game
Date	date of game
Place	place where game was played
Event Name	name of game event

game properties such as the player names and ranks, game date and event name. Player rankings in Go are categorized as follows: professional players are ranked between 1-dan and 9-dan, where the number is increasing in strength; amateur players are similarly ranked 1-amateur dan to 9-amateur dan. Since players are never demoted after achieving a ranking, ranks provide a biased estimate of player skill, particularly for older players. Instead, I rely on calculated skill rankings scraped from Go Ratings (Coulom, 2008), which provide Elo estimates of player skill. Elo rating measures playing skill since it is a mathematical construct that formulates the expected outcome of a game as the Elo difference between both players. Lastly, I complement the information available in the SGF files with individual characteristics sourced from player biographies that accompany the GoGoD database. Table 1 lists variables derived from SGF metadata, Go Ratings, and player biographies that are used in the analysis section.

3.3 Go AI

Recent advances in computer science led to the development of Go-playing AIs with beyond human playing skill. My analysis uses Leela Zero, a reinforcement learning-based open source AI modeled after Alpha Go Zero (Silver et al., 2017), Google DeepMind's Go program that defeated in 2017 the highest ranked player at the time. Leela Zero learns the game exclusively through self-play, i.e. without learning from human games, which means its assessments of move quality are independent of the players performing the move. Despite never receiving human game inputs, Leela Zero's playing style bears resemblance to how humans approach the game in terms of stone grouping, territorial influence, use of forcing moves and other key game concepts. Professionals and amateurs currently use Leela Zero's evaluations as a training tool.

3.4 Construction of Cognitive Performance Measures

I use Leela Zero to measure players' cognitive performance by evaluating the quality of their moves. Mechanically, Leela Zero takes as inputs the setup of pieces on the board, a sequence of the last seven moves played (i.e. a recent game history) and a parameter ρ specifying the number of playouts. Roughly speaking, Leela Zero plays out the remainder of the game ρ times, each time drawing a candidate move from a prior probability distribution given the game history and each time recording whether the game is eventually won or loss. When a position is simple and there is an obvious move choice, Leela Zero's priors leads most of the playouts to attempt the same move. Conversely, complex positions may reflect on many moves with high probability mass, and the ρ playouts will improve the AIs ability to discern which move leads to a higher probability of eventually winning. After Leela Zero finishes evaluating the position, it chooses from the attempted moves the one with highest winning probability.[9] I analyze moves in the game data using $\rho = 2000$ playouts, a number slightly higher than the $\rho = 1600$ used when training Leela Zero through self-play.

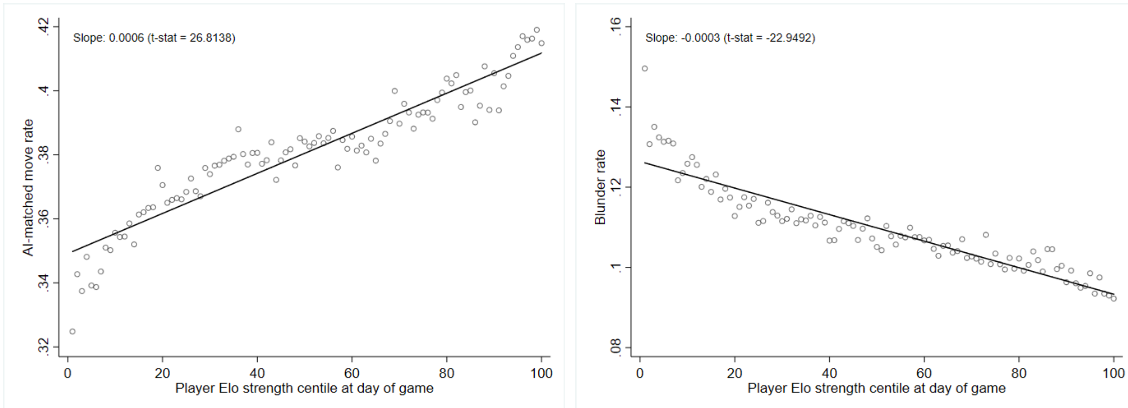
For each move, I construct two binary variables capturing a dimension of cognitive performance. One variable, denoted *AI-matched move*, equals one if the human move in the game data matches Leela Zero's top recommendation. The other, denoted *Blunder*, equals one if the human move is outside the set of candidate moves Leela Zero attempted in its 2000 playouts. These two variables represents extremes on the move cognitive performance spectrum: matching the AI move represents our best estimate of an optimal move, one that requires a high level of reasoning. Meanwhile a blunder represents a significant loss relative to best play, arguably associated with attention lapses.

A sensible check on the validity of our two cognitive performance measures is their relationship with empirical estimates of player skill. Figure 1 plots the averaged outcome variables against player Elo centiles. Averaging the outcome variables gives us the probability that players at a given Elo centile make an AI-matched move or a blunder. Panel A shows a strong increasing relationship between Elo (i.e. player skill) and the probability of AI-matched moves while panel B shows blunders decrease as players become more skilled. Using fitted values, the players in the top 1% skill level make on average 20% more AI-matched moves and 23% fewer blunders than the bottom 1% of the distribution.

In the analysis, I explore the possibility that the effect of air pollution on cognitive performance may be more salient on certain stages of the game. Following Go theory knowledge, I

[9] In reality, the winning probability is a combination of both Leela Zero's priors and the realized winning rates during the playouts, a heuristic similar to Bayesian updating.

Figure 1: Relationship between cognitive performance measures and empirical skill measure



Rates on the vertical axis are the mean of the outcome variables computed at each Elo centile over all moves in the sample. The solid line plots a simple linear regression of each outcome against Elo centile.

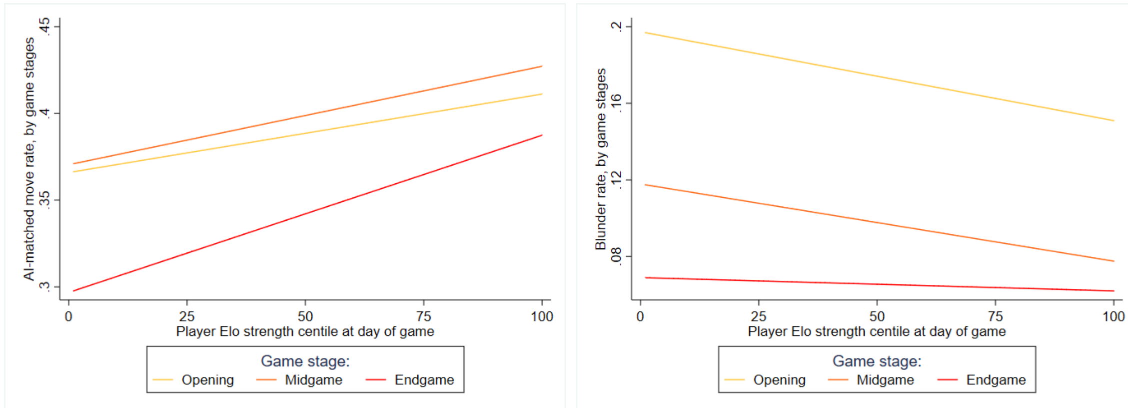
denote the first sixty moves[10] as the opening stage, known by players as *fuseki*. The opening typically involve less systematic play, with moves offering fewer opportunities for contact with the opponent's stones. I denote moves 61 to 160 as the midgame, a stage where systematic play occur more frequently. Stone placements in the midgame offer plenty of opportunities for territorial disputes between players yet the board still offers a large number of plausible move choices. Lastly, I denote moves 161 to the end as the endgame stage. The endgame is characterized by large territories in the board already being established and players pursuing simultaneous disputes over small territorial expansions through the entire board.

Figure 2 shows the regression fit of cognitive performance on player skill for the subsamples of opening, midgame, and endgame moves. For most part, the narrative from figure 1 applies to each game stage separately. The exception to this is the regression fit for endgame blunders in panel B, which shows no relationship between blunders and player skill in that game stage. One possible explanation is that the characterization of blunders is not as meaningful in the endgame. In fact, Leela Zero outputs a much larger number of candidate moves in the endgame, many of which achieving a similar eventual win rate. While there is still a best move in the endgame, it is unclear that there are many move choices that blunder the game.[11]

[10] I refer to a move as one turn taken by one player, which is typically called a ply in the two-player sequential games literature.

[11] Appendix ?? explores how alternative move thresholds for each game stage affect the estimated coefficients in the main specification.

Figure 2: Cognitive performance measures vs empirical skill measure, by game stages



Rates on the vertical axis are the mean of the outcome variables computed at each Elo centile over all moves in the sample. The solid line plots a simple linear regression of each outcome against Elo centile.

4 Data

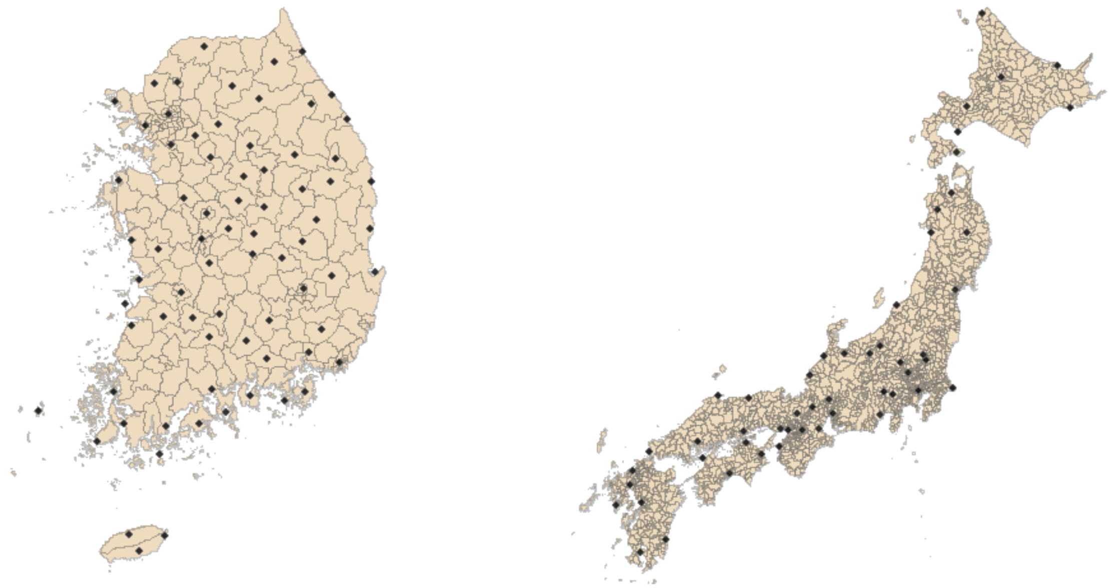
This research explores the relationship between air quality and cognitive performance of Go players. The dataset used combines information on the incidence of the Asian dust in South Korea and Japan, air pollution data in the same two countries, individual characteristics of professional and skilled amateur Go players, and measures of player cognitive performance on each game. The time frame is 1980-2018 and the analysis relies on daily variation of Asian dust across regions located near 28 meteorological stations in South Korea and 53 in Japan. The data pertaining air quality and game records are separately discussed below.

4.1 Air Quality Data

Incidence of Asian dust events and within-dust storm concentration of particulate matter have been on the rise for the last few decades. For many years, meteorological agencies in Korea and Japan have been concerned with the association of air pollution with these dust storms. Both countries adopt a similar strategy for mitigating the health effects associated with the phenomenon by publicly issuing an "Asian dust storm day" warning when meteorological stations detect high concentrations of particulate matter that can be apportioned to dust storms in the desert regions of Mongolia and China.

Records of *hwangsa* (as the dust is known to Koreans) days are available at the Korean Meteorological Administration (KMA) website with daily periodicity from 81 weather stations, spanning the beginning of year 1961 to present day. Similarly, the incidence of Kosa (as its known to the Japanese) is available daily from 59 weather stations at the Japan Meteorological Agency's (JMA) website starting from 1967. Figure 3 presents a map of the

Figure 3: Asian Dust Stations in South Korea (left) and Japan (right)



Circles denote centroid of city where stations are located. All weather stations providing dust incidence data are depicted in the maps. Not all, however, are used in the analysis since many cities in both countries have no game data available.

weather station locations in both countries. For each city, I assign the dust records from its closest weather station by computing Euclidian distances between station and city centroids. I complement the dataset with daily averages of the concentration of PM_{10} and O_3 (South Korea only), SPM and $PM_{2.5}$ (Japan only), SO_2, and CO across 147 monitoring stations in South Korea available from 2001 to 2017 and 218 stations in Japan, available from 2009 to 2016 by the National Institute of Environmental Research (NIER).

4.2 Game Data

The dataset used in the analysis consists of 22,213 games for which all variables listed in Table 1 are available, with at least 120 recorded moves, played between 1980 and 2018 in either South Korea or Japan. Males represent 84% of the players in the sample and play 95% of the games; the age at game date distribution ranges 11 to 97 years old players, with median and mean at 30 and 33 years respectively. 9-dan players amount to 47% of the individuals, and these percentages decrease monotonically as rank decreases.

The game data is matched by date and location with a variable indicating the occurrence of an Asian dust storm at the date and city where the game takes place. As shown in table 2, over 60% of the game records belong to major Korean and Japanese tournaments with individual prize money between $60,000 and $400,000. Major tournaments are annual events and can span over a year from qualifying stages to the finals. The scheduling of games take place long before the actual game date, and many of these events are broadcast live on a regular basis.[12] This is reassuring as it mitigates a concern from Altindag et al. (2017) that dust warnings issued by public authorities lead to avoidance behaviour. While the game rules are consistent across the dataset, Go tournaments may differ substantially in time control systems. Common systems envisage a main period for each player (e.g. 30 minutes or an hour) followed by an overtime protocol, the most common being the so-called *byo-yomi* where players in overtime have a few seconds per move for the remainder of the game.

The key variable for the analysis is a measure of cognitive performance constructed for each player in each game using an AI as expert evaluator. In every game, Leela Zero AI parses the subset of mid-game nodes ranging move 100-119 and stores for each node all candidate moves proposed by its policy network.[13] Early-game moves are too unsystematic for the AI to produce meaningful performance measures. A possible extension for this work would be parsing end-game moves and exploring whether the time constraint after players enter *byo-yomi* alter the estimated effects.

Upon completing the policy network step, each candidate move receives a VN score between zero and one by the AI's MCTS algorithm, where the move strength is increasing in its VN score [14]. I then construct for both game-players: (1) the percentage of "strong moves," where a player's move is defined as strong if it coincides with the move achieving highest VN score in the node; and, (2) the percentage of "blunders," i.e. moves played that are not in the set of candidate moves proposed by Leela Zero. Given that Leela Zero's current playing strength slightly exceeds the strongest Go professionals, these two metrics likely relate with game idioms *tesuji* (strong move) and *poka* (blunder move) respectively. In the analysis, the outcome variables constructed from (1) and (2) are denoted the game-player *strong_percent* and *blunder_percent* respectively.

A remark: Leela Zero proposes a variable number of moves per node, with this number ranging 1 (in special cases such as *sente* forcing moves by the opponent) to 26. Figure 4 plots the histogram of candidate moves per node in my parsed output of move evaluations by Leela Zero.

[12]South Korea and Japan have cable television channels dedicated to Go news and game broadcasts (BadukTV and K-Baduk in Korea, Igo-shogi in Japan). In addition, public TV channels such as NHK (and historically TV Tokyo) are known for offering live coverage on big title matches.

[13]The average number of moves in the sample is 214, which makes this a sensible mid-game range choice to capture strong and blunder moves that occur during *joseki* patterns.

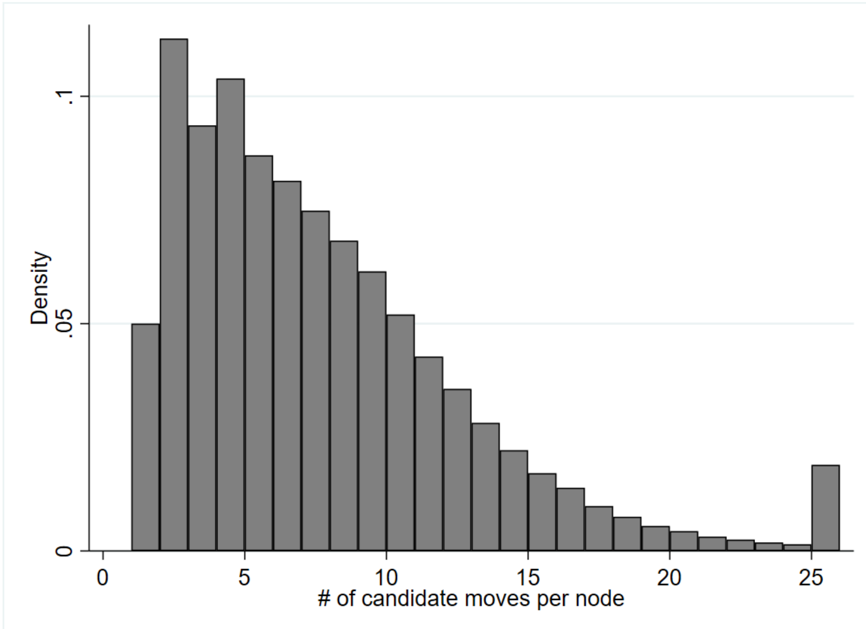
[14]Recall that the VN score represents the win rate of a candidate move in the Monte Carlo simulations.

Table 2: Summary of Major Tournaments in the Data

	tournaments	games	avg duration	% high dan	Prize(USD)
Bacchus	36	328	364	69	unknown
Fujitsu	26	591	224	92	130,000
Gosei	41	1,139	366	97	70,000
GS Caltex	15	273	166	79	60,000
Honinbo	86	1,516	311	89	280,000
Judan	40	1,145	476	97	130,000
Kisei	60	1,382	393	87	400,000
Kiseong	25	290	382	72	unknown
Kuksu	61	473	157	67	unknown
LG	24	607	241	78	60,000
Meijin	79	1,635	350	93	300,000
Myeongin	53	598	201	72	90,000
NEC	37	226	211	98	unknown
Nongshim†	19	256	182	80	440,000
Oza	42	791	425	95	120,000
Paedal	9	80	158	72	unknown
Paewang	26	240	199	81	unknown
Samsung	23	734	151	82	175,000
Siptan	9	266	136	68	unknown
Taewang	15	145	258	77	unknown
Tengen	45	1,246	419	96	125,000
Tong Yang	11	162	235	90	unknown
	782 (Sum)	14,123 (Sum)	273 (Mean)	83 (Mean)	

†: Nongshim cup is a team tournament with five members from each participating country, and the listed prize is for the entire team.

Figure 4: Histogram of candidate moves per node proposed by Leela Zero



Sample mean is approximately 7 candidate moves per node. Excess mass at 26 moves is due to truncation from the upper bound of candidate moves preset by Leela Zero.

5 Empirical Strategy

I investigate the impact of exogenous pollution shocks induced by Asian dust storms on the cognitive performance of Go players. Due to the nature of these dust storms, my analysis sheds light on short-term impacts of an extreme weather event which is becoming increasingly frequent, while abstracting from any long-lasting effects of cumulative exposure. As I focus on a short-term effect, the analysis is kept at the most granular level of observation – the move level – where we make use of the daily regional variation of dust incidence. Cognitive performance is measured by the AI's judgment on the move quality, i.e. the binary variables for AI-matched move and for blunder. The main specification of the analysis is a Linear Probability Model (LPM) which estimates whether the probability of making an AI-matched move or blundering is affected by an Asian dust "treatment" after controlling for potential confounders:

$$Y_{mpjt} = \alpha + \delta Dust_{jt} + X'_{mpt}\gamma + \psi_j + \eta_{year(t)} + \mu_{month(t)} + \phi_p + \varepsilon_{pjt} \tag{1}$$

where Y_{mpjt} indicates either if the move matches the AI recommendation or if the move is a blunder; $Dust_{jt}$ indicates a dust event on city j and day t.

X_{mpt} is a vector of move and player characteristics which may affect the propensity to make an AI-matched move or a blunder. Player controls include demographics such as an age quadratic and female indicator [this seems out of place here; w/o I can call time-varying controls], and player skill measured by Elo rating percentiles on day t.[15] The move controls are the number of moves proposed by the AI in the position and indicators for whether the move is in the opening, midgame, or endgame range. Both of the move controls account for complexity in the position. For example, a position is likely straightforward if the AI recommends only one move option after its search, and complex if the AI ponders and recommends ten alternatives.

Lastly, ψ_j are city fixed effects absorbing the impact of geographical and geological features that correlate with Asian dust incidence. $\eta_{year(t)}$ and $\mu_{month(t)}$ are year and month time trends. ϕ_t absorbs player fixed effects. The error term ε_{mpjt} is clustered at the city level, which allows for correlation within the treatment unit of cities.

The coefficient of interest δ uncovers the effect of an Asian dust shock on the cognitive performance measures of individuals. This effect has a causal interpretation if incidence of Asian dust is as good as random after controlling for player and move characteristics as well as city and time fixed effects.

6 Results

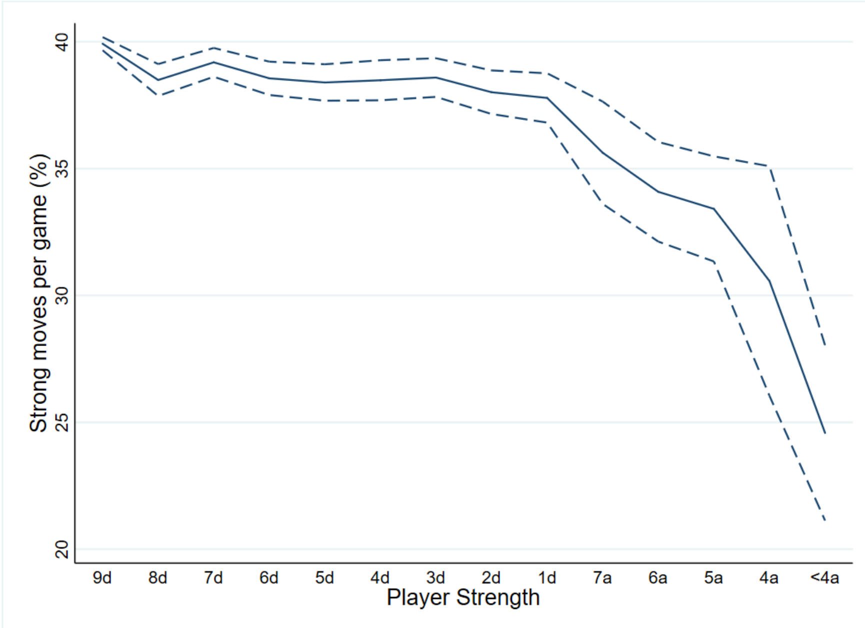
I first examine whether the outcome variables are positively associated with player strength. Figure 5 plots the mean of *strong_percent* across each player rank, ordered from strongest rank at "9d" to weakest at "<4a." The graph shows that the strongest (9-dan) players are also the most likely to play Leela Zero's best suggested move, with the percentage of strong moves per game decreasing about 2 percentage points from strongest to weakest professional dan. This downward relationship kinks and becomes much more pronounced for amateur dan players, which is reasonable since there is a large difference in average ability between the weakest professional players and the strongest amateur players. It is also reassuring that the confidence intervals for these estimates are rather small, such that the difference between professional and amateur rankings is very significant.

Similarly, figure 6 plots the relationship between the mean of *blunder_percent* across player ranks, and the same intuition from Figure 5 holds: players become more likely to blunder as their dan ranking weakens. The strength of these relationships differ between the two metrics: a typical 1-dan player makes 5.3% fewer strong moves but 13.95% more blunders per game than a 9-dan player.[16]

[15]The female covariate is only present in specifications without player fixed effects.

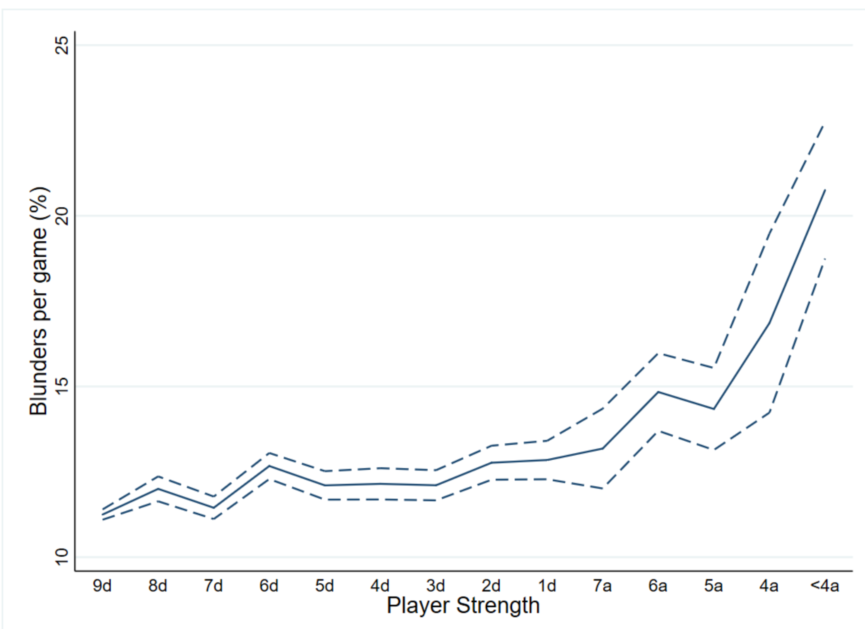
[16]in absolute terms, these percentage changes translate to the weaker professional making 4 fewer strong

Figure 5: Mean percentage of strong moves per game across different ranks



Ranks in x-axis are decreasing in strength from left to right. Percentage of strong moves is calculated is the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the range of moves 100-119.

Figure 6: Mean percentage of blunders per game across different ranks



Ranks in x-axis are decreasing in strength from left to right. Percentage of blunder moves is the percentage of a player's move choices which are not in the set of all candidate moves proposed by Leela Zero.

Since Go is a sequential two-player game, relative cognitive performance of players matter more than absolute performance in determining who wins and loses. If the percentage of strong moves and blunders are in fact measuring the quality of player decisions, making more strong moves or fewer blunders relative to the opponent should increase one's probability of winning. I use the logistic model below to corroborate this claim.

$$F\big[Pr(Black\ wins)\big] = \beta_0 + \beta_1 \mathbb{1}(\Delta strong > 0) + \beta_2 \mathbb{1}(\Delta rank < 0) + \beta_3 \mathbb{1}(\Delta age > 10)\} \tag{2}$$

$$F\big[Pr(Black\ wins)\big] = \beta_0 + \beta_1 \mathbb{1}(\Delta blunder < 0) + \beta_2 \mathbb{1}(\Delta rank < 0) + \beta_3 \mathbb{1}(\Delta age > 10)\} \tag{3}$$

where F is the logit function linearizing the model.

$$F[x] = \ln\left[\frac{x}{1-x}\right] \tag{4}$$

In equation 2, $\mathbb{1}(\Delta strong > 0)$ equals one if the count of strong moves of Black player is higher than White's; similarly in equation 3 $\mathbb{1}(\Delta blunder < 0)$ equals one if Black makes fewer blunders than White. $\mathbb{1}(\Delta rank < 0)$ takes value one if Black's rank is stronger than White's (player strength decreases as the rank number increase); and, $\mathbb{1}(\Delta age > 10)$ equals one if Black is at least ten years older than White.

Columns (1) to (3) of Table 3 present logistic regression results for the model in equation 2, where the coefficients are the exponents of β_1, β_2 and β_3 and can be interpreted as the ratio of winning probabilities when the predictor takes values one versus zero. Columns (4) to (6) of the same table present the results pertaining the model in equation 3. Both tables substantiate the claim that relative performance matters: the odds of winning are approximately 1.5 for the player making more strong moves per game, and the winning odds coefficient is fairly stable to inclusion of other covariates which may predict both the winning outcome and the cognitive performance measures. Similarly, the winning odds are approximately 2.6, with or without additional covariates, for players making fewer blunders. In plain English, these odds ratios translate to players performing relatively more strong moves winning 3 out of 5 games played, and players making fewer blunders winning 8 out of 11 games. Jointly, these table results and figures 5 and 6 validate the move evaluations of

moves and 3 more blunders on a typical game lasting 200 moves.

Table 3: Logistic Regression of Relative Performance on Game Outcome

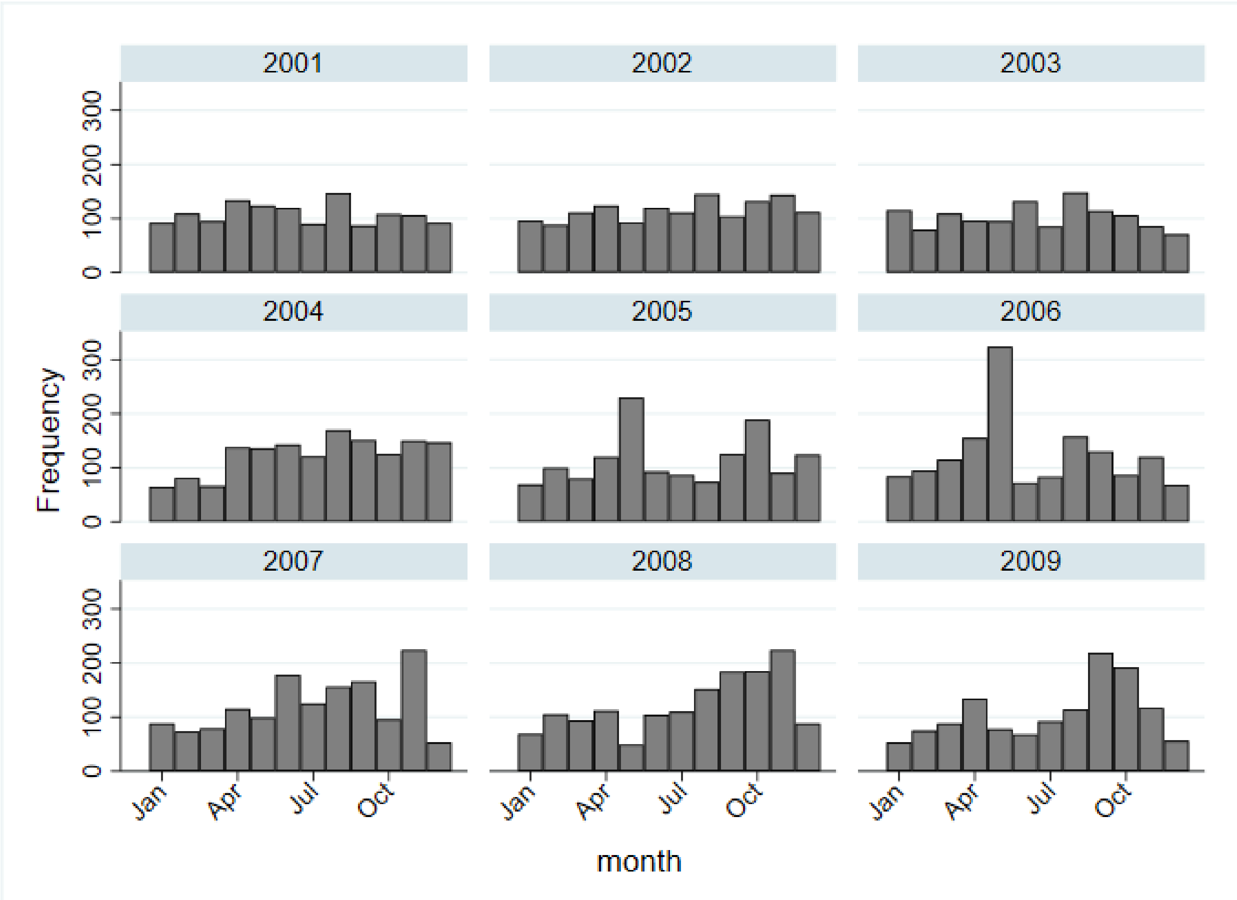
Dep. Var: Pr(Black wins)	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta strong > 0$	1.504***	1.507***	1.510***			
	(14.70)	(14.65)	(14.68)			
$\Delta blunder > 0$				2.677***	2.624***	2.613***
				(35.40)	(34.46)	(34.20)
$\Delta rank$		1.101***	1.124***		1.093***	1.114***
		(18.48)	(21.13)		(16.61)	(19.16)
$\Delta age > 10$			0.659***			0.670***
			(-12.74)			(-11.95)
Observations	22165	22165	22165	22165	22165	22165

Exponentiated coefficients; t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

$\Delta strong$ ($\Delta blunder$) is the difference in count of strong (blunder) moves played by the Black versus White player. Similarly, $\Delta rank$ is the rank difference, where rank is decreasing in strength (rank 2 is weaker than rank 1), and Δage is the age difference of the Black versus White player. The coefficients shown represent an odds ratio, i.e. the probability of winning a game if the predictor is equal to one divided by the probability if it is equal to zero.

Figure 7: Histogram of games played each month over the years ranging 2001 and 2009



Leela Zero as relevant measures of the quality of a player's decision-making in the game.

As mentioned in the previous section, regressing the main specification in equation 2 only yields a causal effect under the assumption that the Asian dust "treatment" $Dust_{jt}$ is conditionally independent of the error term ε_{pjt}. One concern is that some subpopulation of Go players may present a different dust avoidance behaviour than the remainder of the playing population. I investigate this possibility by plotting histograms of games played each month during over a set of years where Asian dust received considerable media attention (see figure 7). While the histogram patterns over time are idiosyncratic, they suggest games are, if anything, more likely to be played during Asian dust season.

While this is not conclusive evidence, it suggests that potential exposure to Asian dust is not a concern taken into account in scheduling games. Most major tournaments observed in the data are events recurring on the same time of the year, with a history that precedes recent concerns about health effects of Asian. History provides anecdotal evidence of unflexible game schedules: the second game of the 3rd Honinbo tournament took place in Hiroshima in August 6, 1945. After the atomic bombing, players and tournament organizers relocated to the city outskirts, where the game was concluded on that same evening.

Evidence against dust avoidance at the player level can be obtained by regressing the following variation of equation 1:

$$Y_{jt} = \alpha + \delta Dust_{j,t+\tau} + \psi_j + \eta_{ym(t)} + \varepsilon_{jt} \quad (5)$$

where Y_{jt} is the log number of games played in day t and city j, and τ represents a window of days around the dust storm. If Asian dust storms are driving a behavioural response where certain individuals forfeit matches, we should observe a decrease in games played on the day of a dust storm (i.e,$\tau = 0$) relative to preceding and subsequent days.[17] Figure 8 presents the estimated $\hat{\delta}$ coefficients and standard errors for leads and lags τ around dust days, where $\tau = -5$ subsumes all days outside the Asian dust event window. While the coefficients are unstable (likely due to sample size limitations), the observed pattern is not consistent with dust avoidance at the player level.[18]

Next, I document the air pollution shock induced by Asian dust storms. An event-study analysis (figure 9) shows a sharp and short-lived increase in daily average concentration of PM pollutants (left panels), between 45% for $PM_{2.5}$ and 75% for PM_{10}, during the day of an Asian dust. For other pollutants (right panels), pollution levels increase modestly at a dust day, but the pictures are somewhat peculiar and involve pollution levels lower than baseline on the day after – for SO_2 and CO – or a few days later in the case of O_3. The decrease in some pollutant concentrations following dust days occurs because higher wind speeds during Asian dust events help eliminate local build-up of man-made pollution that occurs when the air is stagnate (Yang et al., 2017).

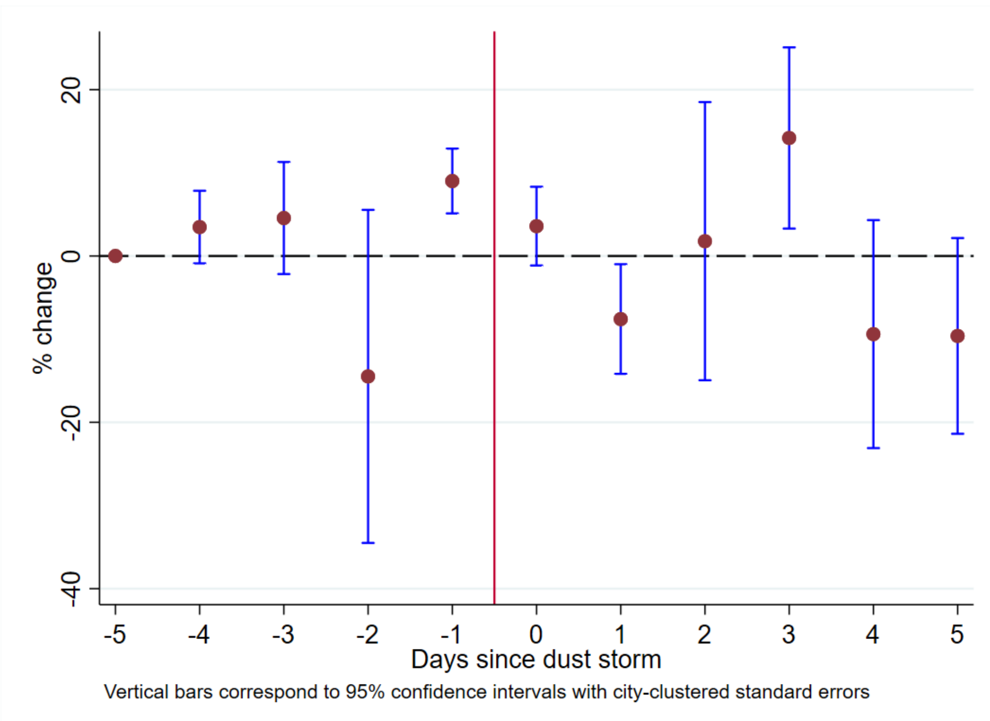
The importance of these shocks also differ across pollutants. A 75% increase in PM_{10} in major South Korean cities translates to an increase of $43\mu g/m^3$ in pollutant concentration[19] – almost sufficient for shifting San Francisco's PM_{10} cleanest recent annual average of $22.8\mu g/m^3$ in 2016 to its dirtiest of $68.8\mu g/m^3$ in 2001. For $PM_{2.5}$ the shock induces an additional $15\mu g/m^3$ increase in pollutant concentration. The shocks diminish in importance for other pollutants. For example, the 8% increase in O_3 during dust days in figure 9 would imply a 1.5 parts per billion (ppb) concentration increase in major South Korean cities, much

[17] This regression focuses on the subset of city-day combinations where games are played, so what I estimate is the effect of Asian dust on the number of games played *conditional* on a game being played in that city and day.

[18] It is nonetheless possible that individuals avoid Asian dusts in other ways, such as wearing masks or choosing different modes of transportation. One should think of the results presented later as being net of such behavioural responses.

[19] Many examples from the health literature – some already cited in this paper – use a $10\mu g/m^3$ change in PM_{10} when reporting adverse effects of this type of pollution.

Figure 8: Change in number of games around dust event window



The y-axis % change is $100 * \hat{\delta}$, i.e. the percentage change in game counts τ days since a dust event. $\tau = -5$ subsumes all game days outside of 5-day event windows.

smaller than the 9ppb gap from dirtiest to cleanest year in San Francisco and also modest in comparison to standards discussed in the health literature. Overall, this event-study analysis points to PM pollution as the main driver of the equation 2 results presented below.[20]

6.1 Overall impacts

Table 4: LPM estimates for effect of Asian dust on cognitive performance

Dependant Variable:	Blunder move				AI-matched move			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dust event	0.003**	0.003***	0.003**	0.002*	0.004*	0.001	0.001	0.001
	(2.87)	(3.60)	(2.94)	(2.52)	(2.32)	(0.44)	(0.49)	(0.54)
Age at game day	-0.000**	-0.000***	-0.001***	0.010	-0.000	0.000	0.000	0.002
	(-3.31)	(-3.49)	(-8.07)	(1.54)	(-0.72)	(1.81)	(0.66)	(0.14)
Age at game day (sq)	0.000*	0.000*	0.000***	-0.000	-0.000	-0.000	-0.000	0.000***
	(2.37)	(2.32)	(6.42)	(-1.87)	(-0.15)	(-0.61)	(-0.21)	(3.69)
Female	-0.001	-0.002	-0.001		-0.006***	0.001	0.002	
	(-0.63)	(-1.39)	(-0.94)		(-7.37)	(0.44)	(1.06)	
Midgame move	-0.078***	-0.086***	-0.086***	-0.086***	0.012***	0.064***	0.064***	0.064***
	(-36.27)	(-41.87)	(-41.92)	(-41.61)	(7.25)	(52.66)	(52.77)	(52.87)
Endgame move	-0.107***	-0.126***	-0.126***	-0.126***	-0.043***	0.086***	0.086***	0.086***
	(-38.35)	(-43.21)	(-43.61)	(-41.99)	(-23.04)	(59.71)	(60.37)	(59.49)
# of moves suggested by AI		0.005***	0.005***	0.005***		-0.036***	-0.036***	-0.036***
		(61.43)	(63.05)	(58.37)		(-184.77)	(-190.09)	(-184.83)
Player strength dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	No	No	Yes	Yes
City FE	No	No	Yes	Yes	No	No	Yes	Yes
Player FE	No	No	No	Yes	No	No	No	Yes
Observations	2931135	2931135	2931135	2931135	2931135	2931135	2931135	2931135
R^2	0.019	0.028	0.028	0.028	0.003	0.170	0.170	0.170

t statistics in parentheses

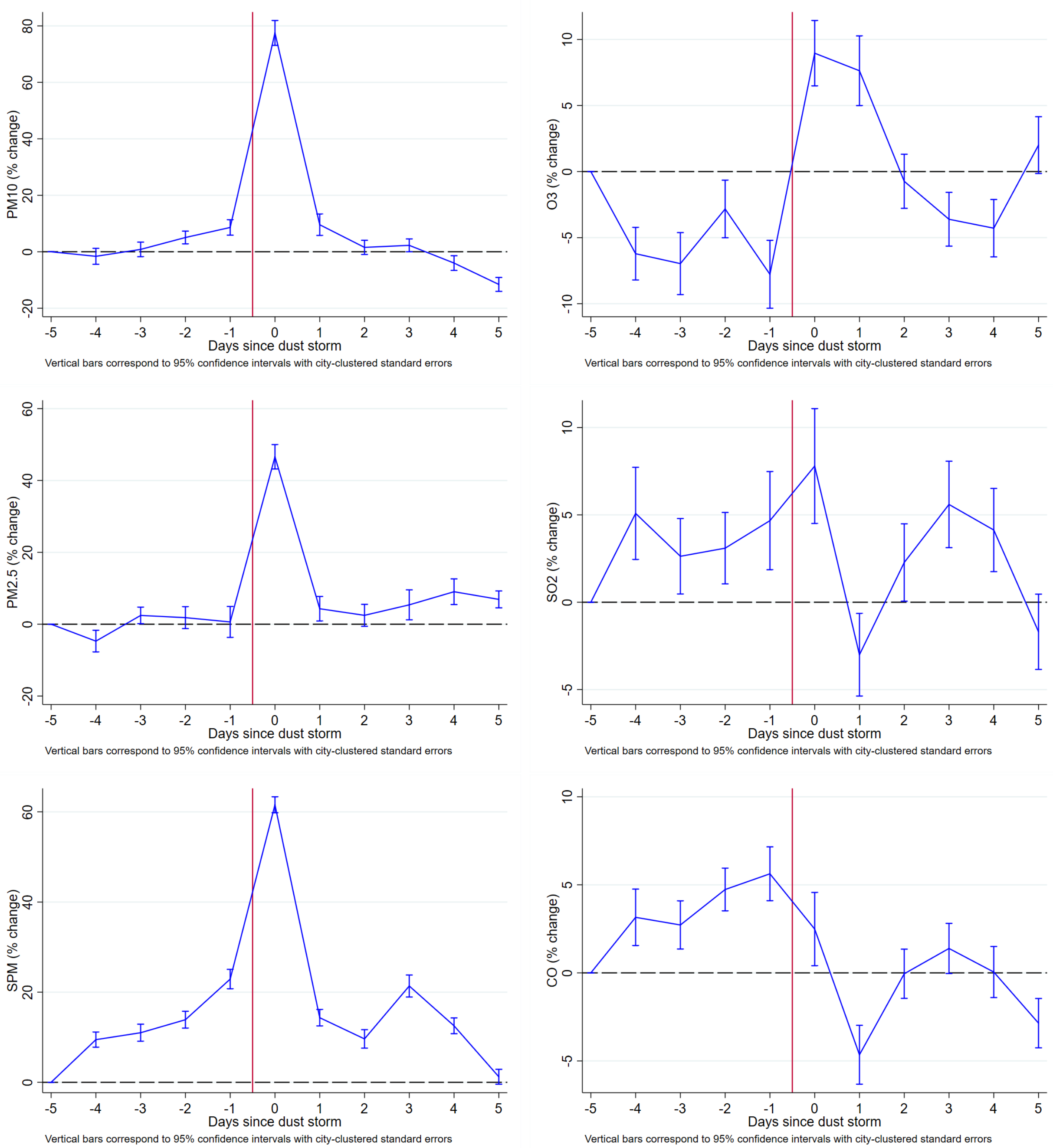
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. Opening moves are the omitted category for game stage dummies.

The results up to this point support interpreting equation 1 as estimating a *causal* relationship of air pollution on the quality of decision-making of Go players. Tables 6- 7 report the estimated impact of Asian dust storms on the constructed measures of Go players' congnitive performance. In all tables, column (1) excludes the control for high, low, and amateur dan as well as city and player fixed effects, column (2) excludes dan controls and player FE, column (3) excludes only player FE, and column (4) estimates the full (and preferred) specification.

[20]San Francisco particulate matter numbers are sourced from the US Environmental Protection Agency Air Trends on Cities and Counties, available online at `https://www.epa.gov/air-trends/air-quality-cities-and-counties`.

Figure 9: Event studies for pollutant levels on five day window before/after dust events



Pollutants examined are PM_{10} (top left), O_3 (top right), $PM_{2.5}$ (mid left), SO_2 (mid right), SPM (bottom left), and CO (bottom right). Data comes from all weather stations in South Korea (years 2001-2017) and Japan (years 2009-2016). PM_{10} and $O3$ observed in South Korea only; SPM and $PM_{2.5}$ observed in Japan only.

Table 5: LPM estimates for effect of Asian dust on cognitive performance, game stage interactions

Dependant Variable:	Blunder move				AI-matched move			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T=1 × Opening move	-0.005	-0.004	-0.005	-0.005	0.003	0.000	0.000	0.001
	(-1.54)	(-1.45)	(-1.52)	(-1.62)	(0.94)	(0.13)	(0.15)	(0.22)
T=1 × Midgame move	0.008***	0.008***	0.007***	0.007***	0.001	0.001	0.001	0.001
	(7.85)	(8.42)	(9.01)	(7.16)	(0.71)	(0.21)	(0.22)	(0.27)
T=1 × Endgame move	0.003	0.004	0.003	0.002	0.007	0.002	0.002	0.002
	(1.14)	(1.67)	(1.84)	(1.54)	(1.70)	(0.77)	(0.76)	(0.85)
Age at game day	-0.000**	-0.000***	-0.001***	0.010	-0.000	0.000	0.000	0.002
	(-3.31)	(-3.49)	(-8.09)	(1.54)	(-0.72)	(1.81)	(0.66)	(0.14)
Age at game day (sq)	0.000*	0.000*	0.000***	-0.000	-0.000	-0.000	-0.000	0.000***
	(2.37)	(2.32)	(6.43)	(-1.87)	(-0.15)	(-0.61)	(-0.21)	(3.68)
Female	-0.001	-0.002	-0.001		-0.006***	0.001	0.002	
	(-0.64)	(-1.39)	(-0.94)		(-7.36)	(0.44)	(1.06)	
Midgame move	-0.079***	-0.086***	-0.086***	-0.086***	0.012***	0.064***	0.064***	0.064***
	(-37.14)	(-43.01)	(-43.07)	(-42.73)	(7.20)	(51.45)	(51.55)	(51.66)
Endgame move	-0.108***	-0.126***	-0.126***	-0.126***	-0.043***	0.086***	0.086***	0.086***
	(-38.22)	(-43.18)	(-43.58)	(-41.93)	(-22.22)	(60.19)	(60.91)	(59.73)
# of moves suggested by AI		0.005***	0.005***	0.005***		-0.036***	-0.036***	-0.036***
		(61.49)	(63.12)	(58.43)		(-184.73)	(-190.05)	(-184.79)
Player strength dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	Yes	Yes	No	No	Yes	Yes
City FE	No	No	Yes	Yes	No	No	Yes	Yes
Player FE	No	No	No	Yes	No	No	No	Yes
Observations	2931135	2931135	2931135	2931135	2931135	2931135	2931135	2931135
R^2	0.019	0.028	0.028	0.028	0.003	0.170	0.170	0.170

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. Opening moves are the omitted category for game stage dummies.

Table 6 reports the estimation results with the percentage of strong moves per game as dependent variable. I find negative yet insignificant point estimates across all specifications, which would suggest that the quality of decision making is not significantly affected by Asian dust exposure. Table 7 reports the results using the percentage of blunders per game as regressand, in which case the estimates are strongly suggestive that Asian dust is associated with an increase in blunders. The coefficients are reasonably stable across specification, and imply that Go players make between 1 and 1.2 percentage point more blunders when exposed to Asian dust. These translate to an increase in blunders between 6.5% and 8.6% relative to days without dust events.[21] Tables 6 and 7 together are suggestive that the overall ability of Go players to inductively reason may not deteriorate, however the likelihood of making a human error increase due to the shock in air pollution exposure.

[21] In absolute terms, the more conservative 6.5% estimates implies two additional blunders in an average-lasting game with 200 moves.

Table 6: Effect of Asian dust on percent of strong moves per game

Dep. Var: Strong moves per game (%)	(1)	(2)	(3)	(4)
Dust event	-0.219	-0.173	-0.166	-0.229
	(0.471)	(0.533)	(0.543)	(0.675)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Dan control	No	No	Yes	Yes
Player FE	No	No	No	Yes
Observations	43755	43755	43755	43755
R^2	0.016	0.023	0.024	0.056

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving highest VN score among the candidate moves evaluated by the AI.

Table 7: Effect of Asian dust on percent of blunders per game

Dep. Var: Blunder moves per game (%)	(1)	(2)	(3)	(4)
Dust event	1.227***	1.235***	1.233**	1.041*
	(0.362)	(0.361)	(0.373)	(0.457)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Dan control	No	No	Yes	Yes
Player FE	No	No	No	Yes
Observations	43755	43755	43755	43755
R^2	0.013	0.017	0.018	0.065

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as a blunder if it is not in the set of candidate moves proposed by Leela Zero's neural network.

6.2 Impacts across age groups

Table 8 reports the regression results by age group using the percentage of strong moves per game as dependent variable. Panel A displays the results for the subpopulation of players younger than 30 years old (the in-sample median age), and has the same flavour as the table 6 results. The coefficients are again statistically insignificant, with precisely estimated point estimates close to zero. The estimates in Panel B – which presents the estimation for players above median age – became more negative however remain insignificant at the 10% significance level. The divergence in coefficient magnitudes in table 8 is slightly indicative that some heterogeneity may exist across age groups.

Table 9 likewise reports the results by age group, using instead the percentage of blunders per game as the regressand. Once again the point estimates for players aged 7 to 30 years old are small and insignificant. However, players between 30 and 96 years old make between 1.8pp and 2.2pp more blunders per game on average, and this estimate is significant at either the 5% or 1% level. These estimates imply an increase in human error between 14.7% and 18.2% for the older players. Compared to Table 7, these results are strongly indicative of heterogeneity, with older players being more susceptible to deterioration of cognitive performance on dust days. This output corroborate public health literature findings of heterogeneous health effects due to Asian dust exposure in South Korea.

6.3 Impacts across player ranks

Figures 5 and 6 have shown a strong relationship between the cognitive performance outcomes and player strength measured by Dan levels. In light of this, I investigate whether the cognitive performance of players differing in strength responds differently to the pollution shocks induced by Asian dust storm. Because high-dan players are better represented in the game records, I combine low-dan and amateur-dan players for this part of the analysis (dust events affecting amateur dan players are particularly rare in this dataset). Table 10 reports the output from regression equation 1 using the percentage of strong moves outcome. Perhaps surprisingly, panel A shows coefficient estimates for the "weaker" players becomes positive (although far from significant) in some specifications, which would suggest air pollution improves the quality of decisions for these individuals. The coefficient returns to being negative (albeit fairly close to zero) on the preferred specification which includes player fixed effects. The output of panel B is very similar to the output for the overall population which is mostly represented by these individuals. The coefficients are negative, statistically insignificant, and fairly stable across specifications, which as before suggest little relationship between air quality and this margin of decision-making quality.

Table 8: Effect of Asian dust on percent of strong moves per game by age group

Dep. Var: Strong moves per game (%)	(1)	(2)	(3)	(4)
Panel A				
Below median age (30 yrs)				
Dust event	-0.130	0.025	0.006	-0.260
	(0.558)	(0.652)	(0.673)	(0.866)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Dan control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.030	0.040	0.041	0.080
Observations	21427	21427	21427	21427
	(1)	(2)	(3)	(4)
Panel B				
Above median age (30 yrs)				
Dust event	-0.952	-1.191	-1.155	-1.207
	(0.835)	(0.845)	(0.836)	(0.825)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Dan control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.027	0.037	0.038	0.077
Observations	21427	21427	21427	21427

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving highest VN score among the candidate moves evaluated by the AI.

Table 9: Effect of Asian dust on percent of blunders per game by age group

Dep. Var: Blunder moves per game (%)	(1)	(2)	(3)	(4)
Panel B				
Above median age (30 yrs)				
Dust event	2.204***	2.185***	2.153***	1.836**
	(0.547)	(0.554)	(0.555)	(0.643)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Dan control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.024	0.032	0.033	0.093
Observations	21427	21427	21427	21427

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as a blunder if it is not in the set of candidate moves proposed by Leela Zero's neural network.

Table 11 reports the estimates using the percentage of blunders outcome as dependent variable. The effects of air pollution on blunders is visible and significant at the 5% level for both stronger and weaker players. Low-dan and amateur-dan players suffer the largest loss in cognitive performance, with blunder moves increasing 2.15pp (an additional 4 blunders per average-duration game). Blunders by high-dan players increase 1pp, implying on average two more blunders per game. A possible yet highly speculative explanation for these differing results is as follows. While professional players in all dan levels exhibit determination and dedicate a large amount of time to improving their game performance, the individuals who make it to the top may have certain traits – such as constancy and higher concentration – turning them more "resistant" to health shocks induced by air pollution.

Go professionals, with the exception of a handful at the very top, earn modest wages for tournament participation and eventual tournament wins. The weaker professionals in particular have to complement their earnings either by becoming teachers or by securing side jobs. These individuals are, if anything, more representative than the high-dan counterpart of individuals working on mental tasks requiring similar cognitive functions as displayed by Go players.

6.4 Robustness

Lastly, I reproduce the results from the preferred specification (column 4) for all tables above after dropping from the sample players who did not play any games (recorded in the data) during Asian dust days. The age at game date distribution for the remaining players is comparable to when using the full sample. The rank at game date however is now even more clustered at the highest ranks, which is consistent with the story of games at the highest levels conforming to a fixed, unalterable schedule in spite of poor air quality conditions. The output from this robustness check, presented in table 12, shows coefficient estimates as well as patterns of statistical significance which are fairly consistent with the results from previous subsections.

7 Conclusion

I have exploited regional and time variation in the incidence of meteorological phenomena known as Asian dust to establish a relationship between air pollution and quality of decision-making of high-level players of the board game Go. I first document that Asian dust storms induce an air pollution shock which raise short-term concentration of coarse and fine particulate matter by 75% and 45%, and also causes some modest increases in other

Table 10: Effect of Asian dust on percent of strong moves per game by player strength

Dep. Var: Strong moves per game (%)	(1)	(2)	(3)	(4)
Panel A				
Low/Amateur-Dan				
Dust event	0.841	0.570	0.678	-0.092
	(0.939)	(0.746)	(0.673)	(0.737)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Age control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.061	0.071	0.076	0.182
Observations	8540	8540	8178	8178
	(1)	(2)	(3)	(4)
Panel B				
High Dan				
Dust event	-0.332	-0.324	-0.260	-0.340
	(0.626)	(0.669)	(0.688)	(0.783)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Age control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.018	0.026	0.026	0.053
Observations	35215	35215	34686	34686

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving highest VN score among the candidate moves evaluated by the AI.

Table 11: Effect of Asian dust on percent of blunders per game by player strength

Dep. Var: Blunder moves per game (%)	(1)	(2)	(3)	(4)
Panel A				
Low/Amateur-Dan				
Dust event	1.369	1.717*	1.714*	2.150**
	(0.967)	(0.769)	(0.749)	(0.720)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Age control	No	No	Yes	Yes
Player FE	No	No	No	Yes
R^2	0.056	0.063	0.068	0.189
Observations	8540	8540	8178	8178
	(1)	(2)	(3)	(4)
Panel B				
High Dan				
Dust event	1.233***	1.259***	1.159***	1.008**
	(0.260)	(0.270)	(0.264)	(0.351)
Female=1	Yes	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
City FE	No	Yes	Yes	Yes
Age control	No	No	Yes	No
Player FE	No	No	No	Yes
R^2	0.015	0.020	0.021	0.061
Observations	35215	35215	34686	34686

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Years covered in the regression are 1980 to 2017. A player's move is defined as a blunder if it is not in the set of candidate moves proposed by Leela Zero's neural network.

Table 12: Robustness check: excluding players with zero Asian dust treated days

Sample:	(1) Full	(2) Younger	(3) Older	(4) Low/Amateur-Dan	(5) High-Dan
Strong (%)	-0.373	-0.861	-1.105	-0.677	-0.492
	(0.699)	(0.776)	(0.896)	(0.591)	(0.845)
Blunder (%)	0.931*	0.801	1.501*	2.681***	0.739**
	(0.398)	(0.828)	(0.636)	(0.662)	(0.273)
N	28785	13190	15302	3256	25236

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regressions in these table reproduce the preferred specification for tables 6 to 11 after excluding from the sample the players who did not play any game in an Asian dust day.

pollutants.

I also construct productivity measures for Go players aided by state-of-the-art advances in artificial intelligence. Using an AI which outperforms even the best players of this game, I evaluate moves played by expert Go players and classify them as strong, acceptable, or blunder. I demonstrate that cognitive performance measures constructed from these move evaluations are correlated with player strength, and also that relative performance within game matters for determining the game winner and loser.

My main findings, based on evaluations of games played in South Korea and Japan during Asian dust days versus "clean" days, strongly suggest that the air pollution shock induced by Asian dust causes Go players to blunder 6.5% more. In absolute terms, this amounts to roughly two additional blunders during a standard game lasting 200 moves. These estimated air pollution effects dissipate for players less than 30 years old and become more pronounced for older players, which is consistent with evidence that older individuals are more susceptible to adverse health effects from air pollution. I also find heterogeneous effects by player strength: lower-ranked professionals and amateurs – which I argue are more likely to resemble other decision-makers in the population – make on average four additional blunders per game. For a reference, Archsmith et al. (2018) estimate a remarkably similar effect: Major League of Baseball umpires make one additional ball/strike incorrect call for every 250 decisions when $PM_{2.5}$ levels increase $10\mu g/m^3$.

In contrast, I find no significant evidence across various specifications and multiple player demographics that air pollution affects Go players' ability to make strong moves. The outcomes from this research suggest air pollution may have little to no effect on the quality of good decisions, but also that poor air quality induces an increase in human error. It is hard to tell how *economically* meaningful these results are. Playing Go demands specific cognitive skills and so my findings may have limited validity for a broad population. Previous research make the case that expert chess players – who are perhaps the most similar individuals to expert Go players – are known to use inductive reasoning in their work and exhibit higher than average ability to backward induct. This suggests my results have implication to other decision-makers whose work involves careful consideration and some degree of uncertainty. My analysis does not speak of the underlying channel driving these results, although I point the readers to a health literature linking particulate matter exposure to central nervous system disorders such as migraine and headache which potentially leads to deterioration of cognitive functioning. Understanding the biological link between air quality and human error is left as an exciting avenue for future research.

8 Extensions

Two extensions to this work require no additional data. A third extension requires collecting data for games played online. These ideas are discussed below.

8.1 Cumulative exposure to air pollution

The effects of continued exposure to air pollution is gaining momentum both academically (Carey et al., 2018) and in popular media (Leung, 2019). Multiple dust storms in consecutive days are uncommon, but it may nonetheless be possible to capture the effect of multiple dust days with a different regression specification. For this, the dust dummy $Dust_{jt}$ would be replaced with a count of dust days for each individual in the sample.

8.2 Depth of reasoning

Behavioural economics sustains an interest on situations where human decisions deviate from rational, "optimal" strategies. Since the origins of bounded rationality (Simon, 1957), many experiments have been conducted with the purpose of understanding how deeply individuals reason. Staged contests such as the "Blotto game" and "p-beauty game" provide us with estimates of game theory concepts in bounded rationality, such as k-level thinking.

Strong moves and blunders are two extreme cases on the measure of quality of decisions in Go. Can move choices in general reveal more about the thought process of players? Work in algorithmic decision theory by Biswas and Regan (2015) have successfully measured chess players' skill from assessing the quality of their moves. One extension to my research is applying their work to estimate the *depth* of Go moves.

Table 13: Shimako vs Hirohisa 7th Shinjin-O Leela Zero win rate values at various depths

Move	Depth → 1	2	3	4	5	swing
G2	10.32	11.15	12.25	11.15	12.60	5.96
N11	10.32	12.75	13.22	12.89	12.49	1.21
J14	10.30	10.30	10.30	10.30	11.21	4.07
O11	11.97	12.17	11.50	11.53	10.88	-3.22
L13	10.02	12.35	11.37	11.48	10.02	-4.71
J13	7.60	7.60	7.60	8.14	7.60	-0.11

Moves are represented as alphanumeric coordinates in a 19×19 grid. Numbers in depths $d = 1$ to 5 are the move values, i.e. the probability of eventually winning if move is played.

Let $\delta_d(m)$ denote the difference in value from move m to the best move when we are constrained to a certain thinking depth d. This depth d can be related to thinking time and the human's speed of information processing. $\delta_d(m)$ is then the difference from optimality of move m at depth d. Define move swing as the sum of differences from optimality from each depth d and the highest depth D:

$$sw(m) = \sum_{d=1}^{D} \delta_d(m) - \delta_D(m) \tag{6}$$

Table 13 provides an example of the move swing output in a particular board configuration of a played game from the dataset. Moves with a swing value greater than zero (colored blue) become more attractive at greater thinking depths, while moves with negative swing (colored red) become less attractive instead.

Figures 10 and 11 plot the average error at each depth of all moves played by high-ranked and low-ranked professionals, respectively.[22] The intersection between the average error of human moves and of the AI moves represents the highest depth the human considered prior to choosing a move.

In principle, the d calculated in these figures can be constructed for all players in the data and used as an outcome variable of cognitive performance. Work remains to be done, however, in interpreting the connection between changes in d and boundedly rational decision-making.

[22] That is, $\frac{1}{N_m}\sum_m \delta_d(m)$ where the summation is over all moves played by high-ranked or low-ranked players. N_m denotes the number of moves in the set being summed.

Figure 10: Average error of swingdown moves, low-ranked players and Leela Zero

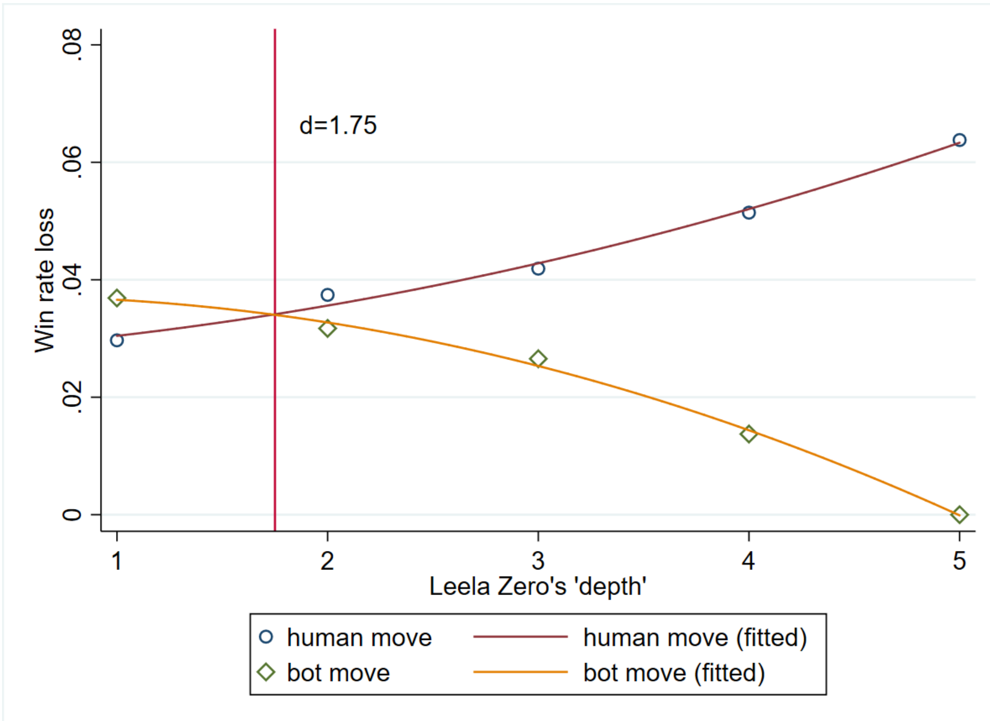
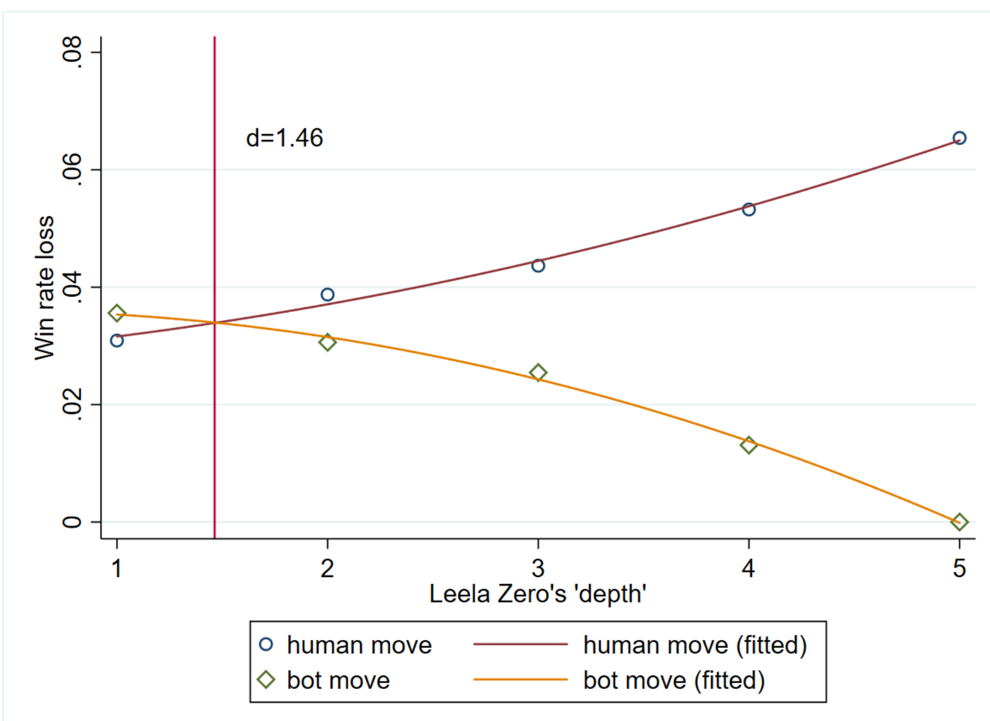


Figure 11: Average error of swingdown moves, low-ranked players and Leela Zero



8.3 Air pollution changes game outcomes

A feature of the games that make into the GoGoD database is that they are typically a match-up of individuals with similar skill level, but also within the same age group, and sex. If some individuals are more susceptible to the adverse effects of air pollution, air pollution becomes a determinant of game outcome.

I have collected data from 1.5 million online games played between 2005 and 2016 and currently am working on identifying player locations on the day the online games are played. Possibly, the matching of two players in this online platform is less systematic and offers more variation in skill differences, age gaps, and matching of players from different sex. Rather than parsing these games with an AI, I will explore how a player's probability of winning is affected by Asian dust storms. Possible scenarios are: (1) the player is exposed to Asian dust and the adversary is not; (2) only the adversary is exposed to Asian dust; and, (3) both are exposed but the players differ in age or some other characteristic linked to heterogeneous effects in the literature.

References

Altindag, D. T., Baek, D., and Mocan, N. (2017). Chinese yellow dust and korean infant health. *Social Science & Medicine*, 186:78–86.

Archsmith, J., Heyes, A., and Saberian, S. (2018). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. *Journal of the Association of Environmental and Resource Economists*, 5(4):827–863.

Backus, P., Cubel, M., Guid, M., Sánchez-Pages, S., and Mañas, E. (2016). Gender, competition and performance: Evidence from real tournaments. *IEB Working Paper*.

Biswas, T. and Regan, K. (2015). Measuring level-k reasoning, satisficing, and human error in game-play data. In *Machine Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on*, pages 941–947. IEEE.

Carey, I. M., Anderson, H. R., Atkinson, R. W., Beevers, S. D., Cook, D. G., Strachan, D. P., Dajnak, D., Gulliver, J., and Kelly, F. J. (2018). Are noise and air pollution related to the incidence of dementia? a cohort study in london, england. *BMJ open*, 8(9):e022404.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., and Mullainathan, S. (2016). Productivity and selection of human capital with machine learning. *American Economic Review*, 106(5):124–27.

Chay, K. Y. and Greenstone, M. (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The quarterly journal of economics*, 118(3):1121–1167.

Chun, Y., Cho, H.-K., Chung, H.-S., and Lee, M. (2008). Historical records of asian dust events (hwangsa) in korea. *Bulletin of the American Meteorological Society*, 89(6):823–828.

Coulom, R. (2008). Whole-history rating: A bayesian rating system for players of time-varying strength. In *International Conference on Computers and Games*, pages 113–124. Springer.

Currie, J., Zivin, J. G., Mullins, J., and Neidell, M. (2014). What do we know about short-and long-term effects of early-life exposure to pollution? *Annu. Rev. Resour. Econ.*, 6(1):217–247.

French, H. (2002). China's growing deserts are suffocating korea.

Genc, S., Zadeoglulari, Z., Fuss, S. H., and Genc, K. (2012). The adverse effects of air pollution on the nervous system. *Journal of Toxicology*, 2012.

Ghio, A. J., Kim, C., and Devlin, R. B. (2000). Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers. *American journal of respiratory and critical care medicine*, 162(3):981–988.

Gobet, F., Retschitzki, J., and de Voogt, A. (2004). *Moves in mind: The psychology of board games*. Psychology Press.

Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–73.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3):689–730.

Hausman, J. A., Ostro, B. D., and Wise, D. A. (1984). Air pollution and lost work.

Herskovitz, J. (2008). China's killer yellow dust hits korea and japan.

Jia, R. and Ku, H. (2015). Is chinas pollution the culprit for the choking of south korea? evidence from the asian dust. Technical report, Working Paper.

Keene, R. and Levy, D. (1992). *How to beat your chess computer*. Henry Holt and Co., Inc.

Kwon, H.-J., Cho, S.-H., Chun, Y., Lagarde, F., and Pershagen, G. (2002). Effects of the asian dust events on daily mortality in seoul, korea. *Environmental research*, 90(1):1–5.

Lavy, V., Ebenstein, A., and Roth, S. (2014). The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. Technical report, National Bureau of Economic Research.

Lee, H., Kim, H., Honda, Y., Lim, Y.-H., and Yi, S. (2013). Effect of asian dust storms on daily mortality in seven metropolitan cities of korea. *Atmospheric environment*, 79:510–517.

Lee, J.-T., Son, J.-Y., and Cho, Y.-S. (2007). A comparison of mortality related to urban air particles between periods with asian dust days and without asian dust days in seoul, korea, 2000–2004. *Environmental research*, 105(3):409–413.

Leung, E. (2019). Pollution could be damaging your brain, research suggests. *The Globe and Mail*.

Levitt, S. D., List, J. A., and Sadoff, S. E. (2011). Checkmate: Exploring backward induction among chess players. *American Economic Review*, 101(2):975–90.

Loane, C., Pilinis, C., Lekkas, T. D., and Politis, M. (2013). Ambient particulate matter and its potential neurological consequences. *Reviews in the neurosciences*, 24(3):323–335.

Mori, I., Nishikawa, M., Tanimura, T., and Quan, H. (2003). Change in size distribution and chemical composition of kosa (asian dust) aerosol during long-range transport. *Atmospheric Environment*, 37(30):4253–4263.

Mosteller, D. (2016). Air pollution's hazy future in south korea.

Ozkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E., and Jenkins, P. (1996). Personal exposure to airborne particles and metals: results from the particle team study in riverside, california. *Journal of Exposure Analysis and Environmental Epidemiology*, 6(1):57–78.

Palacios-Huerta, I. and Volij, O. (2009). Field centipedes. *American Economic Review*, 99(4):1619–35.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017). Mastering the game of go without human knowledge. *Nature*, 550(7676):354.

Simon, H. A. (1957). Models of man; social and rational. *Oxford press*.

Yang, Y., Russell, L. M., Lou, S., Liao, H., Guo, J., Liu, Y., Singh, B., and Ghan, S. J. (2017). Dust-wind interactions can intensify aerosol pollution over eastern china. *Nature communications*, 8:15333.