2023-08

# Essays on Productivity and Innovation

Novaes de Amorim, Arthur

UNIVERSITY OF CALGARY

# Essays on Productivity and Innovation

by

Arthur Novaes de Amorim

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ECONOMICS

CALGARY, ALBERTA

AUGUST, 2023

# Abstract

Productivity and innovation are essential components for economic growth. This dissertation studies the determinants of productivity and innovation from various perspectives. The first chapter investigates the historical role taken by mass media in increasing agricultural productivity. I use the early twentieth-century establishment of commercial radio in the United States to quantify the impact of locally relevant farm programming on productivity growth. Using variation in exposure to radio due to topography, my analysis shows that the broadcasting of local farm programming led to an increase in the productivity of land used in agriculture that persisted for at least two decades. This positive effect was not limited to a certain region, and was felt in a variety of important crops grown across the country. Consistent with radio reducing information barriers, the productivity gains were more pronounced for farmers in areas with lower literacy rates and economic status, lower media saturation, and reduced transport connectivity via railroads.

While the first chapter is about productivity growth in the scale of a country, the second chapter zooms into the productivity of an individual high-skilled worker. I provide evidence that Asian dust storms affect the cognitive performance of high-skilled individuals as they make complex decisions in the strategy board game Go. I develop a novel data set linking historical records of high-level Go games with localized measurements of dust storm activity. Using a powerful artificial intelligence as an expert evaluator of over 400,000 game moves, I examine how quasi-random variation in exposure to Asian dust events affect player performance. I document that dust storms lead to a short-lived but sharp increase of on average $75 \mu g/m^3$ in $PM_{10}$. My main results show players exposed to Asian dust on the game day remain able to find the best moves in a position, but also become more susceptible to human error, making 8.3% more inaccurate moves. I subsequently establish that these

adverse effects on human error are mostly driven by older individuals while players younger than 30 years old are not significantly affected by the deteriorated air quality. My findings reveal a hidden cost of air quality for mature workers performing tasks that require mental acuity and involve critical thinking, satisficing, and other problem-solving concepts demanded in various modern professional occupations.

The final chapter studies, in the context of energy storage, the interplay between market size, policy, and innovation. I utilize geolocated data on energy storage projects in the United States to show that the arrival of a new project in a county causes an immediate increase in local patenting activity on storage-related technologies. This effect is short-lived, contrary to what might be expected if the project results in local knowledge spillovers that culminate in follow-on innovation. I additionally examine a policy change that increases the potential market size for energy storage solutions by compensating frequency regulation resources for increasing regulation capacity, accuracy, and speed. This policy led to a significant increase in the national output of energy storage-related patents. These findings suggest that incentivizing the expansion of energy storage projects can indirectly help promote innovation, eventually allowing energy storage to realize its full potential in supporting the transition to clean energy generation.

# Preface

This dissertation is original, unpublished, independent work by the author, Arthur Novaes de Amorim.

# Acknowledgments

I am deeply grateful to my supervisor, Alex Whalley, for his unending (but never excessive), thoughtful advice and ever-present guidance and attentiveness throughout the development of this thesis. I am especially grateful for his generosity with his time. I would like to thank my committee members, Lucija Muehlenbachs and Jean William Laliberté, for their encouragement and helpful comments whenever I was struggling with research. You were tremendous role models.

Beyond my committee, numerous faculty members helped me grow during the Ph.D. journey. Special thanks to Shawn Kantor at Florida State University, for countless interactions in our collaborative work. I was fortunate to receive great instruction in my early years of grad school, and I thank in particular Arvind Magesan and Scott Taylor. Outside of economics, Vineet Saini and Rob Deardon instilled in me a deep interest in interdisciplinary work.

I am grateful for others who made those years in Calgary worth remembering, especially Cosmas Dery, Nahiyan Azad, Abraham Saucedo Cepeda, Jing Jiang, Ghaz Malekitabrizi, Mokhtar Tabari, Chris Michaud, Hugo Tran, and Clarence Kim.

A heartfelt thank you to my parents and siblings for encouraging me to embark on this journey. Lastly, thank you to my loving spouse, Hazel, and my children, Luah, Yule, and Gio, for their unconditional love and support. I could not have made it without you. Thank you.

Dedicated to Hazel, Luah, Yule, and Gio.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Agricultural Change in the United States: Evidence from the Golden Age of Radio

## 1.1 Introduction

One of the biggest questions in economics is: why cross-country differences in productivity are so large? It is a well-known and persistent empirical fact that poor countries have large agricultural sectors with disproportionately low productivity when compared to rich countries (Gollin et al., 2014). One often cited explanation for this fact is the slow adoption of new technologies in developing countries, possibly in part due to limited access to information. Because the agricultural sector thrives on information that is specific to geography, climate, and other local factors, understanding the role of different sources of information can offer important policy implications.

This paper investigates the role of mass media on the provision of localized information that enhances agricultural production. Many studies suggest learning and information frictions as key determinants of agricultural productivity (Foster and Rosenzweig, 1995; Conley and Udry, 2010). Yet, empirical work unpacking the effect of mass media on agricultural productivity is limited. Here, I examine the role of radio broadcasting on agricultural change in America during the 1920s to 1950s. This period in time, known as the "Golden Age" of radio, offers a unique setting as it coincides with a dramatic transformation in agriculture with the developments of high yield crop varieties, chemical fertilizers, soil conservation practices, and innovations in farm machinery. I analyze the impact of exposure to local farm educational programming from radio on agricultural productivity on the short and medium run and seek to understand the mechanisms through which this information channel influenced agricultural growth.

I compile a novel data set of digitized records with technical information on all commercial AM radio stations in operation between the 1920s to 1950. I also gather from historical sources a list of educational radio stations with an emphasis on broadcasting locally relevant farm programming. Using an engineering model of sound propagation, these data allow me to predict across space and over time the degree of a county's exposure to farm radio, proxied by radio signal strength. I pair the radio data with county-level panel data from the U.S. agricultural census covering this time period (Haines et al., 2014), measuring various agricultural outcomes such as the value of agricultural land, the aggregate value of crops, and the production of major cash crops.

Distinguishing the informational effects of radio from other amenity effects is a difficult task. One econometric challenge concerns the endogenous location of radio stations. As an urban phenomenon, commercial radio stations face a problem of maximizing advertising revenue which is tantamount to maximizing listenership in cities.[1] On the other hand, educational radio stations broadcasting local farm programming are often associated with and co-locating at universities. I address this endogenous location concern by exploiting spatial exogenous variation in signal strength. The identification comes from (1) the opening and closing of radio stations over time, (2) changes over time in broadcasting technology resulting typically in increased radiated power from a station's transmitters, and (3) an empirical strategy first used by Olken (2009) to exploit residual spatial variation in the strength of AM radio signals due to topographic factors.

Another challenge concerns the bundling of all other forms of radio programming not related to the agricultural sector, but which may nonetheless impact the livelihood of agricultural workers. I explore this issue by examining the effects of exposure to farm-focused radio stations versus all other (non-farm-focused) radio stations in an attempt to isolate the effect of the provision via radio of information that is relevant to local farmers.

[1]While half of American urban homes had a receiver by 1930, only 27 percent of rural homes did. (Craig, 2006)

I start by documenting that counties with higher exposure to farm radio displayed higher overall agricultural productivity. Specifically, a one standard deviation increase in the signal strength of farm radio increased the per acre value of farm land by 2.1%. I find suggestive evidence that this effect was larger in counties with less access to alternative sources of information through other radio stations, in counties with decreased transportation connectivity measured by proximity to railroads, and in counties with lower literacy rates and lower economic status measured by averaged occupational income scores.

To ensure that these first findings did not conflate productivity with other amenity effects of radio that could influence land prices, I show the effect is robust to quantifying agricultural productivity with a revenue-based measure of the per acre value of all crops combined. Using this alternative productivity measure, a one standard deviation increase in farm radio signal led to a 4.4% increase in crop value per acre. I also conduct a falsification test using exposure to other radio stations that did not place an emphasis on farm content. This test shows that the main results were not driven by any radio exposure *per se*, suggesting instead the effects were unique to farm radio programming. I finally unpack the effects of farm radio on the productivity of five of the largest cash crops grown across the entire U.S. during the time. I find significant positive results for all but one crop, ranging from 3.9% in oat yields to almost 10% on cotton yields with a one standard deviation increase in the signal strength of farm radio.

In a related empirical strategy, I exploit the residual variation over time and space in signal strength to estimate the dynamic effects of farm radio with an event study design. I find the effect on overall agricultural productivity measured by the per acre value of farm land persisted throughout the decades of radio's Golden Age and amounted to approximately 8% over the 1920s to 1940s for the counties that received farm radio signal early on, but the effect on the per acre value of crops is short-lived. Taken together, all the findings show that mass media can lead to persistent growth on the agricultural sector.

This paper contributes to the literature on the effects of information access and learning on agricultural productivity (for survey papers related to this literature, see Aker (2011), Bridle et al. (2020), and Suri and Udry (2022)). In particular, my paper is closely related to recent ongoing work by Gupta et al. (2020), who study the role of mobile phones on technology adoption and productivity in agriculture, though our papers differ in my emphasis on a one-way and affordable form of mass communication through radio. My paper is also related to a variety of social projects and randomized trials conducted in the developing world seeking experimental evidence of the impact of radio on farmers' knowledge and welfare. My work complements this body of research with a historical lens from the perspective on a developed country during a time when alternative sources of information were scarce in rural communities. Lastly, recent work by Kantor and Whalley (2019) emphasize the local nature of spillovers from universities on agricultural productivity in the late nineteenth century. As their estimated spillover effects dissipate within 20 years, their findings suggest a reduction in the value of information diffusion that occurred through interactions in close proximity between farmers and researchers early in the twentieth century, as these interactions were supplanted by new technologies allowing for long distance communication such as the telephone and radio.

I also contribute to a growing body of work employing electromagnetic signal propagation models to study the effects of mass media. Social science researchers have used these models to study mass media's impact on a variety of contexts such as public spending (Strömberg, 2004), social capital (Olken, 2009), and political persuasion (Enikolopov et al., 2011; DellaVigna et al., 2014; Adena et al., 2015; Gagliarducci et al., 2020; Wang, 2021).

## 1.2 Historical Background

### The Expansion of Radio in the US

The origins of broadcasts from commercial radio stations trace back to the start of the 1920s. As a new medium for entertainment and educational information, radio quickly became a household favorite throughout the US. The immediate popularity of radio is evidenced by its rapid expansion. Panel (a) of Figure 1.1 shows the share of American families owning a radio increased steadily from zero to 4o% at the end of the decade and 80% by 1940. The number of commercial stations climbed sharply from zero in 1920 to roughly 600 by the end of the decade. Sales of radio equipment increased fourteenfold during the same time period (panel (b)) and an estimated fourteen million US homes owned radios by 1930.

**(a) Radio penetration** **(b) Sales of radio equipment**

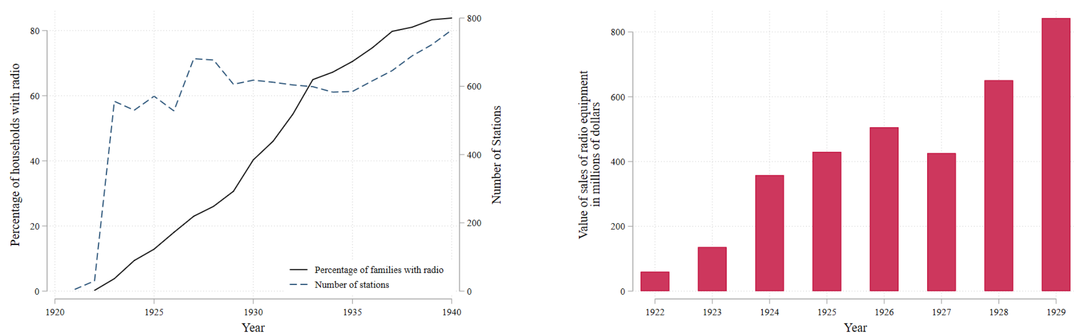



Figure 1.1: Growth of radio, 1920 to 1940

*Notes:* Data for panel (a) is from the 1940 *Broadcasting Yearbook* (Broadcasting Publications, Inc., 1959). Sales data for panel (b) is from (Douglas, 1987).

On a radio conference in 1922, recognizing the value of radio for farmers, then secretary of commerce Herbert Hoover stated that “no single use of radio should take precedence over its use for agriculture...” In fact, farmers knew to tune in at specific stations for weather forecasts and crop reports, as well as educational talks on agricultural technologies (Wik, 1981). The US Department of Agriculture (USDA) was heavily involved in the production of farm programs targeting the dissemination of frontier technologies of relevance to farmers. These included nationwide programs providing general advice to farmers, such as the famous *National Farm and Home Hour*, *Farm Flashes*, and *Housekeeper Chats*. Appendix Exhibit 1 shows selected excerpts of radio programming compiled from transcripts of the *National Farm and Home Hour*.[2]

On a local level, farming information was delivered by state agricultural radio programs associated with land-grant colleges, state universities, and state agricultural extension services. Land grants often had strong agricultural programs and a close relationship with the USDA and state and local agricultural organizations, so the match-up was natural. By the end of 1922, “stations

[2]The transcripts are available at the Internet Archive Collection `https://archive.org/details/usda-nationalfarmhomehour`.

such as the University of Wisconsin's WHA, WOI at Iowa State College, WKAR at Michigan State, and Texas A&M's WTAW were all carrying a regular schedule of locally-produced agricultural broadcasts" (Craig, 2001). The typical educational radio station dedicated approximately one-fifth of airtime to market and technical information for farmers (Tyler, 1933). Radio played an important role in diffusing innovations stemming from research performed at state agricultural experiment stations.

While the societal value of radio for farmers was widely recognized, the preferences of rural audiences were not a focus of most commercial radio stations. Even though rural listenership increased substantially over time, few advertisers were willing to pay to reach an audience they considered to be made up of relatively unimportant consumers. Anecdotally, farmers sought entertainment on the evenings in commercial radio stations but agricultural information during daytime, highlighting the importance of government-sponsored educational radio service.

## Radio and the Modernization of Rural Life

At the beginning of the 20$^{\text{th}}$ century, rural America was being left behind in technological and social evolution. Concerned with the economic consequences of the increasing gap between urban and rural areas, many reformers suggested closing the gap through innovations with the potential of integrating farmers, such as the telephone, the automobile, electricity, and radio. By the 1920s these four technologies had been introduced in rural areas, and by 1930 half of farms in the US had automobiles. The radio boom of the 1920s was largely an urban phenomenon, and adoption at farms was initially slower due to lack of electrification[3] and to the high cost of radio equipment. As the radio industry matured, manufacturers began marketing battery-powered "farm radios" and by 1940 more farms owned radios than had telephones, automobiles, or electricity (Craig, 2006).

[3] Only 10% of farm homes had electricity by 1930.

## 1.3 Data and Empirical Strategy

The empirical strategy seeks to quantify the effect of exposure to local farm content broadcast in AM radio on agricultural productivity and understand how provision of information impacted farmers' productivity decisions.

A key strength of this analysis is the ability to leverage exogenous variation in exposure to radio due to the impact of local topography on the propagation of radio waves.[4] I first calculate for each radio station-county pair the "free space" signal strength where the Earth is assumed to be a smooth surface, devoid of any features which may act as barriers to the propagation of radio waves. The free space signal strength at any point is inversely proportional to the square distance from a radio transmitter and also a function of the electrical conductivity of the ground and the broadcasting technology. Next, using the method developed by Olken (2009), I attenuate the free space signal strength by the propagation loss due to geographical features on the path from the radio transmitter to the receiving county. This is done with an off-the-shelf implementation of the Irregular Terrain Model (ITM), developed by the U.S. government and considered an industry standard for predicting broadcasting signal strength (Oughton et al., 2020). Below, I describe the radio data used to measure point-to-point signal strength of AM radio stations at the centroid of each county.

### Data

This paper utilizes novel data on radio availability during the first half of the 20$^{\text{th}}$ century. Data for estimating the signal strength of radio stations is drawn from multiple sources. From the World Radio History Project, I collect information about all commercial radio stations' transmitter power, antenna height relative to ground level [5], and broadcast frequency starting in 1922. These data are cross-checked in different years through various sources (Radio Age, 1927; Radio Digest,

[4] The path of AM signal propagation varies throughout the day. While at nighttime the AM signal travels long distances through "skywave" propagation, signals in daytime travel by conduction over the surface of the Earth. Importantly, most farm programming occurs during morning hours and at noon, where topography plays a role on how far the signal travels.

[5] Antenna height is known only in 1940 (Broadcasting Publications, Inc., 1959), and when missing it is predicted through a regression of height on the log of the transmitter power for other years. See Appendix Figure A1 for the linear fit of this regression.

1933; USDA, 1933; Radio Annual, 1950; Broadcasting Publications, Inc., 1959) for completeness. Data on ground conductivity – also utilized for the signal strength calculations – comes from the Media Bureau of the Federal Communications Commission. The topographic profile between a transmitting and receiving points comes from a digital elevation model with 1/3 arc-second (10 meters) spatial resolution.[6]

I classify a station as a farm radio station if it is listed in the State Agricultural Radio Programs section of Brunner (1936), compiled in a symposium with inputs from program directors, managers of land-grant college radio stations, heads of agricultural colleges, farm group executives, editors of agricultural publications, and members of State Departments of Agriculture and State Extension Services. Importantly, this classification implies the location of farm radio stations is closely related to the location of land-grant colleges and State-run extension services.[7]

Using these data, I calculate the point-to-point signal strengths between county centroids and the city coordinates of each radio station and assign to each county the maximal signal strength. This operation is done separately for farm radio stations and for other (non-farm) radio stations, resulting in a panel data set measuring county-level radio predicted signal strength on five year intervals ranging from 1925 to 1950.[8]

**Radio Signal Strength.** Figure 1.2 shows the predicted signal strength of farm radio stations– measured in decibel-milliwatts (dBm) – resulting from the free space ($FarmSignalFree$) and irregular terrain models ($FarmSignal$). The dBm metric is commonly used in radio communication to express absolute power levels. It here serves as a proxy for the quality of radio reception within each county. For ease of interpretation in the analysis, signal strength will be expressed in standard deviations from the mean, with one standard deviation in 1925 corresponding to 20.8dBm for farm radio stations and 14.2dBm for other radio stations.

The panel data set measuring county-level radio signal strength is linked with the following

[6] Sourced from the National Elevation Database developed by the U.S. Geological Survey (USGS, 2017).

[7] Historically, the location of these land-grant colleges in the 19$^{th}$ century depended on a variety of political, environmental, and geographical factors. Moretti (2004) supports the idea that "the geographical location of land-grant colleges seems close to random" from the perspective of later developments.

[8] I use radio stations available on the rollout years of agricultural censuses (e.g., 1924 for the 1925 agricultural census). Due to this factor, and availability constraints, the final data draws from published lists of commercial US radio stations in the years 1924, 1929, 1934, 1938, 1945, and 1950.

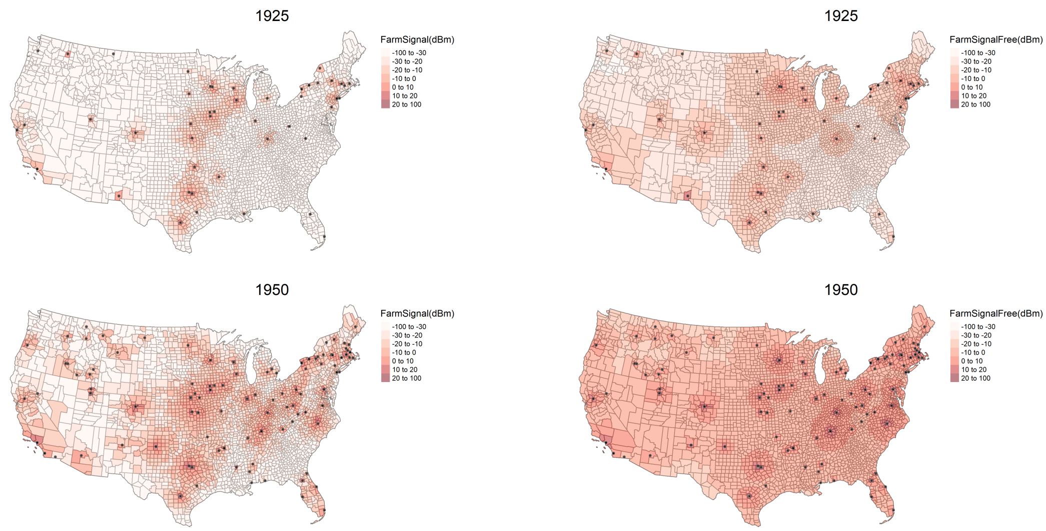



Figure 1.2: Signal strength of radio stations broadcasting farm content

*Notes:* The left figures show the strongest predicted (computed using the ITM) signal strength of farm radio stations in each county, and the right figures similarly show the signal strength in free space, all measured in dBm. Data for 1925 (the top figures) comes from various published lists of radio stations in 1924, and data for 1950 (the bottom figures) from published lists in 1950 as described in the data section.

data: time-invariant county-level environmental data on soil quality from Fishback et al. (2005); gridded terrain elevation (USGS, 2017) from which a county-averaged terrain ruggedness index is constructed following Nunn and Puga (2012); time-invariant gridded crop suitability from the Global Agro-Ecological Zones (Fischer et al., 2021) project of the Food and Agricultural Organization (FAO), from which county-level average suitability indices are computed for various crops; and time-varying gridded historical climate data from the PRISM Climate Group (PRISM, 2011), used to construct annual cumulative precipitation and mean temperature at the county level.

Lastly, the panel data set includes agricultural census records from Haines et al. (2014) and population census records from Haines (2005). These data contain key agricultural and socioeconomic information used in the analysis and are described in the subsection below.

**Census Data.** I obtain historical county-level agriculture panel data by combining all waves of the decennial and quinquennial editions of the census of agriculture between 1910 and 1950. Additional socioeconomic data is included from the population census covering the same time period. Linear interpolation is used for some key agricultural and socioeconomic variables in years in which data is not reported in the census. The first two census periods in the sample (1910 and

1920) predate the first commercial radio station and are used mainly for covariates balance checks assuming a counterfactual distribution of radio stations available in 1925.

**Sample selection.** While I collect census data starting from 1910, the first time period in the analysis is 1925 and the sample selection steps below consider the time periods included in the analysis when restricting the sample. I take the following steps to create the balanced panel of counties used in the analysis. Firstly, I map the historical data from all different years and sources into modern county boundaries utilizing the crosswalk developed by Eckert et al. (2020). Then, I drop counties in the top and bottom 1% (pooling data from 1925 to 1950) of the per acre value of farm land and per harvested acre value of all crops. I also drop counties with reported acres of land in farms that exceeds the county's total land area. These observations are dropped due to measurement error in the agricultural census and measurement error introduced by the weights in the county boundaries crosswalk. Lastly, I drop counties with less than 1,000 acres of land and counties that report less than 20% of land in farms in any census year between 1925 and 1950, which for the most part are highly urbanized counties or regions with a topography unfavourable for farming. The resulting balanced panel comprises 2,230 counties within the continental U.S. with modern-day (2010) boundaries, observed over six agricultural census 5-year periods from 1925 to 1950. Panel (a) of Appendix Figure A4 depicts the counties featured in the "baseline" main sample.

**Descriptive Statistics.** Table 1.1 shows descriptive statistics of key variables for the year of 1925, the year of the first census of agriculture since the establishment of commercial radio stations in the US. Column (1) presents the mean and standard deviation of relevant variables for the baseline full sample of 2,230 counties remaining after sample selection. Columns (2) and (3) present similar statistics for the subsamples of counties above and below the median predicted signal strength of farm-focused radio stations. Column (4) shows the p-value associated with a test for difference in means between the subsamples from columns (2) and (3).

The Statistics presented in Table 1.1 illustrate how the distribution of farm radio exposure was far from random. Counties with above median signal strength in 1925 had significantly higher agricultural productivity (panel A) – as measured by farm and crop value per acre – and were more populated (panel B), having on average 1.3 times the population of those below the median signal strength. Panel B also confirms that counties with above median signal strength had a relatively larger agricultural sector. Panel C shows that, as expected, counties with a farm radio station by 1925 will have higher radio penetration in the near future, as proxied by the percentage of farm families with radio by 1930. The panel also confirms the mechanical relationship between predicted signal strength and fixed county factors such as terrain ruggedness and ground conductivity. The results from this table highlight the importance of a well-designed empirical strategy for dealing with endogeneity of radio from its early days.

## Empirical strategy

To examine the short-run impact of farm radio, I use the following two-way fixed-effects estimation equation:

$$Y_{ct} = \beta_1 FarmSignal_{ct} + \beta_2 FarmSignalFree_{ct} + \delta X_{ct} + \gamma_c + \theta_t + \varepsilon_{ct}. \qquad (1.1)$$

In this equation, the variable $Y_{ct}$ is an agricultural productivity outcome of county $c$ in year $t$, such as farm value per acre. $FarmSignal_{ct}$ represents the maximum predicted signal strength received at the centroid of the county by a radio station broadcasting farm content (hereon "farm radio") in that year and $FarmSignalFree_{ct}$ represents the maximum signal strength assuming unobstructed signal propagation. $X_{ct}$ is a vector of controls for socioeconomic characteristics, climate, and in the richest specification includes an interaction of soil characteristics and year dummies. The baseline specification also includes year ($\theta_t$) and county ($\gamma_c$) fixed effects that absorb national trends and time-invariant characteristics across counties. Errors are corrected for clustering at the county level.[9]

The coefficient of interest is $\beta_1$, measuring the effect of exposure to farm radio on a given agricultural outcome variable. Far from being randomly placed, radio stations typically locate in areas that maximize listenership and consequently advertising revenue. $FarmSignalFree$ alleviates this endogeneity concern as it partials out the decision to locate in densely populated areas, leaving us with residual variation in exposure to farm radio due to topography.

[9] I test the robustness of my main estimates to an alternative method that accounts for spatial correlation in the error terms (Conley, 1999) in Appendix Table A1.

Identification requires this residual variation in signal strength to be unrelated to unaccounted changes in determinants of agricultural productivity. While this assumption requirement cannot be directly tested, Figure 1.3 presents standardized estimated regression coefficients of farm radio predicted signal strength in 1925 on key outcomes, various predictors of agricultural productivity from census data, and crop productivity in 1920, prior to the establishment of commercial radio stations. These estimates essentially allow us to test for effects of farm radio exposure *prior* to exposure, at a time where no effect would be expected.

Without controlling for $SignalFree$, the test shows that the measure of farm radio signal strength does predict agricultural outcomes and relevant demographic characteristics, confirming what was shown using 1925 data in Table 1.1. The predictive power of signal strength is expected since the location of radio stations five years later was not random. Upon accounting for this source of endogeneity by controlling for $SignalFree$, the estimated coefficients become smaller in magnitude and in most cases statistically insignificant at the 5% significance level. More is done for the sake of identification in the main analysis, where I can rely on the advantages of the panel setting by including controls and county and year fixed effects. Nonetheless, the results from Figure 1.3 illustrate the need for controlling for free space radio signal to mitigate threats to identification.

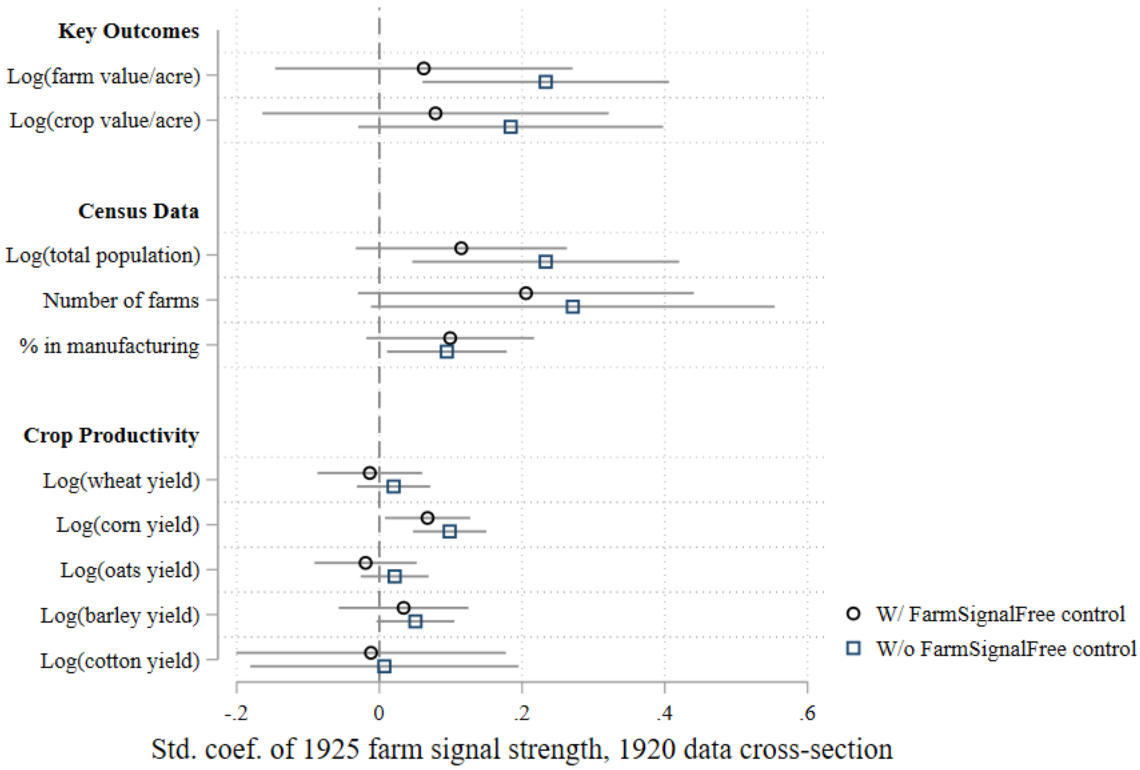



Figure 1.3: Balance tests

*Notes:* Plotted standardized coefficients are for the $FarmSignal_c$ variable in a regression on 1920 cross-sectional data of the form $Y_c = \beta_1 FarmSignal_c \ [+\beta_2 FarmSignalFree] + \delta_s + \varepsilon_c$. Regressions includes state fixed effects ($\delta_s$), and errors are clustered at the state level. The grey lines represent 95% confidence intervals. Predicted farm radio signal strength is a 5-year lead (i.e., the 1925 computed signals), as there were no commercial radio stations by 1920, and regressors are constructed from the 1920 population and agricultural censuses.

Table 1.1: Descriptive statistics, main sample of US counties in 1925

| | (1) All | (2) Above median farm signal | (3) Below median farm signal | (4) Diff. p-value |
|---|---|---|---|---|
| ***Panel A: Key Outcomes*** | | | | |
| Farm value $/acre | 60.83 | 77.14 | 44.52 | 0.00 |
| | (44.19) | (50.01) | (29.57) | |
| Crop value $/acre | 10.20 | 10.47 | 9.93 | 0.03 |
| | (5.963) | (4.941) | (6.825) | |
| ***Panel B: Crop Productivity*** | | | | |
| Wheat yield (bushels/acre) | 15.01 | 17.24 | 12.62 | 0.00 |
| | (6.062) | (6.282) | (4.775) | |
| Corn yield (bushels/acre) | 21.31 | 24.53 | 18.09 | 0.00 |
| | (9.223) | (9.072) | (8.192) | |
| Oats yield (bushels/acre) | 24.78 | 28.93 | 20.57 | 0.00 |
| | (10.38) | (10.18) | (8.763) | |
| Barley yield (bushels/acre) | 21.74 | 23.55 | 18.80 | 0.00 |
| | (8.051) | (7.577) | (7.935) | |
| Cotton yield (bales/acre) | 0.36 | 0.32 | 0.38 | 0.00 |
| | (0.128) | (0.167) | (0.0987) | |
| ***Panel C: Census Data*** | | | | |
| Population (000s) | 32.90 | 37.42 | 28.38 | 0.00 |
| | (55.65) | (68.96) | (37.41) | |
| Number of farms (000s) | 2.40 | 2.45 | 2.34 | 0.06 |
| | (1.332) | (1.201) | (1.450) | |
| % employed in manufacturing | 6.26 | 5.80 | 6.72 | 0.01 |
| | (7.858) | (7.696) | (7.994) | |
| ***Panel D: Radio Penetration*** | | | | |
| % farm families with radio (1930) | 22.21 | 30.63 | 13.78 | 0.00 |
| | (19.81) | (19.86) | (15.77) | |
| Strongest farm radio signal (dBm) | -47.86 | -30.06 | -65.66 | 0.00 |
| | (21.95) | (11.24) | (14.28) | |
| Mean Terrain Ruggedness Index | 41.51 | 35.09 | 47.94 | 0.00 |
| | (50.46) | (39.10) | (59.02) | |
| Ground conductivity | 9.39 | 11.56 | 7.21 | 0.00 |
| | (8.254) | (8.172) | (7.754) | |
| Number of counties | 2230 | 1115 | 1115 | 2230 |

Standard deviations in parentheses.

*Notes:* This table shows the mean of 1925 county characteristics (except for the % of farm families with radio in panel D, which is sourced from 1930 data). Column (1) shows the means over all counties in the main sample as described in the data. Columns (2) and (3) show the means over the subgroups of counties with predicted farm signal strength above and below median in 1925, respectively. Column (4) shows the p-value of a $t$ test for the difference in the means in columns (2) and (3).

## 1.4 Results

**Short-run effects of farm radio.** The first result presented sheds light on how exposure to farm radio affected overall agricultural productivity on the short run. My main outcome of interest for overall productivity is the per acre value of farm land and buildings (hereon "farm value per acre") as it may capture various dimensions in which farm radio can improve agricultural practices of farmers. The drawback of this measure is it may correlate with other determinants of farm land value unrelated to agricultural productivity, and as such it may conflate any utility derived from listening to radio with changes in productivity due to provision of farming content. I address this concern by investigating the effect of farm radio on additional outcomes that provide narrower measurements of farm productivity. Later in the analysis, I also attempt to isolate any possible effect unrelated to productivity by incorporating to the regression model other radio stations which provide arguably the same utility to households, minus the farming content.

The results of estimating equation 1.1 with this outcome are reported in Table 1.2. Moving rightward across the table, we go from sparsest to richest specification. Column (1) contains only county and year fixed effects, and controls are added for farm signal in free space in column (2) and for various socioeconomic and environmental factors in column (3). Column (4), the preferred specification, allows time-invariant county soil characteristics – which may factor on ground propagation of radio waves – to have different marginal effects over time through the inclusion of soil characteristics interacted with year dummy variables. Across all specifications, predicted farm radio signal strength has a positive coefficient significant at better than 1%. Consecutively adding covariates attenuate the estimated coefficients, but do not affect statistical significance. The estimated effect from the preferred specification implies a one standard deviation increase in signal strength leads to 2.1% higher farm value per acre.

The initial results from Table 1.2 could be attributed to channels unrelated to agricultural productivity, such as other utility value derived from radio exposure. I unpack these initial results firstly by examining more directed measures of agricultural productivity, related to crop revenue and crop quantities, as outcome variables. These narrower measures may not be individually relevant to

Table 1.2: Farm Radio Exposure and Farm Value per Acre

| Dependent variable: | Log(farm value/acre) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $FarmSignal$ | 0.049*** | 0.034*** | 0.031*** | 0.021*** |
| | (0.004) | (0.006) | (0.005) | (0.005) |
| County fixed effects | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ |
| $FarmSignalFree$ | | ✓ | ✓ | ✓ |
| Baseline controls | | | ✓ | ✓ |
| Soil characteristics × year controls | | | | ✓ |
| Observations | 13,380 | 13,380 | 13,380 | 13,380 |
| Number of Clusters | 2,230 | 2,230 | 2,230 | 2,230 |
| Adjusted R-Squared | 0.945 | 0.945 | 0.951 | 0.958 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes:* $FarmSignal$ is a standardized measure (mean zero and variance one) of the predicted signal strength of farm radio resulting from the Irregular Terrain Model. $FarmSignalFree$ in contrast is the predicted signal strength assuming a smooth and featureless earth. Baseline controls include log of total population and farm population, percentage of farms with tenancy regime, and percentage of males, Black individuals, and manufacturing workers. Soil characteristics include soil water capacity, % of soil consisting of clay, soil erodibility (K) factor, soil drainage quality, liquid limit of the soil layer, and soil annual flood frequency. Farm value per acre is the ratio of the combined value of all farms and buildings over the acres of farm land. Standard errors are corrected for clustering at the county level.

farm radio listeners in the entire US, as crop mix varies largely across the country. Taken together, they offer a more complete picture of the initial results on farm value per acre.

Table 1.3 presents estimates of the preferred specification of equation 1.1 with these additional outcomes. Column (1) of the table shows the estimated coefficients using the value of all crops combined (hereon "crop value per acre") as dependent variable. This outcome quantifies the revenue productivity of cropland. As a revenue-based measure, the overall value of crops may not only capture increases in crop output, but also price differences due to crop quality or due to price dispersion resulting from frictions in the crops market (Kantor and Whalley, 2019). The estimated effect implies a one standard deviation increase in signal strength leads to 4.4% higher crop value per acre, a result that is quantitatively larger than the overall productivity effect using farm value per acre.

Columns (2) to (6) of Table 1.3 zoom into five of the largest crops grown in the first half of the 20$^{\text{th}}$ century U.S.: wheat, corn, barley, oats, and cotton. These are crops that experienced drastic changes in yields during this time period, both losses due mainly to soil erosion during the Dust Bowl in the 1930s and gains due to technological progress such as advances in plant breeding and chemical fertilizers. With the exception of corn, exposure to farm radio appears to have increased the yield of the examined crops, with the estimated effect ranging between 3.9% for oats and 9.9% for cotton with one standard deviation increase in signal strength. Throughout the country, corn experienced arguably the steepest increase in yield among the examined crops, due partly to developments in corn hybridization. Perhaps word of corn innovations spread fast regardless of radio, as evidenced by the quick adoption of hybrid corn in Griliches (1957) [10].

Table 1.3: Farm Radio Exposure and Agricultural Productivity – Crop Value and Yields

| | Overall crop value | Crop productivity | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Log(crop value/acre) | Log(wheat yield) | Log(corn yield) | Log(barley yield) | Log(oat yield) | Log(cotton yield) |
| $FarmSignal$ | 0.044*** | 0.051*** | -0.002 | 0.062*** | 0.039*** | 0.099*** |
| | (0.009) | (0.008) | (0.008) | (0.012) | (0.008) | (0.011) |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| County fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $FarmSignalFree$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Baseline controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Soil characteristics × year controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 13,380 | 11,859 | 13,257 | 9,872 | 12,907 | 5,109 |
| Number of Clusters | 2,230 | 2,091 | 2,226 | 1,894 | 2,218 | 878 |
| Adjusted R-Squared | 0.829 | 0.615 | 0.792 | 0.485 | 0.597 | 0.684 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* $FarmSignal$ is a standardized measure (mean zero and variance one) of the predicted signal strength of farm radio stations resulting from the Irregular Terrain Model. $FarmSignalFree$ in contrast is the predicted signal strength assuming a smooth and featureless earth. Baseline controls include log of total population and farm population, percentage of farms with tenancy regime, and percentage of males, Black individuals, and manufacturing workers. Soil characteristics are time-invariant, and include soil water capacity, % of soil consisting of clay, soil erodibility (K) factor, soil drainage quality, liquid limit of the soil layer, and soil annual flood frequency. Crop value per acre is the ratio of the aggregate value of all crops over the acres of harvested cropland. Sample size varies by crop due to differences in which counties have a positive planted area for specific crops, as yields are otherwise undefined. Standard errors are corrected for clustering at the county level.

**Event study design.** The evidence presented so far offers insights on the short-run effects of

[10]Estimates in Griliches imply it took between 4 and 12 years for hybrid corn diffusion to go from 10 percent to 90 percent (Manuelli and Seshadri, 2014).

farm radio on agriculture. I now turn to a different specification seeking to understand the dynamic cumulative effect of farm radio on agricultural productivity. I do so with a research design that considers exposure to farm radio as treatment events of identical intensity that occurs at the county level, potentially on multiple time periods, following the estimation notation laid out in Schmidheiny and Siegloch (2023).

Treatment assignment occurs when the farm radio signal strength of a county exceeds a threshold, here defined as the median predicted signal strength at 1925, the first period since the opening of commercial radio stations. With multiple events of identical intensity, this implies that treatment $T_{c,t}$ is a dummy variable equaling 1 in any period where the county's farm radio signal exceeds the threshold, and 0 otherwise. I estimate the following equation of levels on changes in treatment:

$$Y_{c,t} = \sum_{\ell=-3}^{4} \beta_\ell D_{c,t-\ell} + \delta X_{c,t} + \gamma_c + \theta_t + \varepsilon_{c,t}, \tag{1.2}$$

where $Y_{c,t}$, $X_{c,t}$, $\gamma_c$, and $\theta_t$ are defined as in equation 1.1. $D_{c,t-\ell} = \Delta T_{c,t-\ell}$ is the first difference of the treatment status, and this status is assumed to remain constant beyond the endpoints of the event window. The dynamic treatment effects parameters are $[\beta_{-3}, \beta_{-2}, 0, \beta_0, ..., \beta_4]$, normalized to one period prior to the treatment assignment, i.e., $\beta_{-1} = 0$. The periods described in the horizontal axis between each agricultural census are five-year gaps, with the exception of the 1910 to 1920 census, which has a ten year gap.

Figure 1.4 shows the cumulative effect of farm radio on the two main agriculture productivity measures under the dynamic model with slightly different treatment assignment given by equation 1.2 where counties are treated if the ITM-predicted signal strengths exceeds the median signal in 1925. While I find no evidence of a pre-treatment effect for farm value per acre on panel (a), the estimated coefficient on one of the pre periods for crop value per acre in panel (b) is significantly different from zero. To interpret these results, the cumulative effect on farm value per acre would imply counties where farmers were exposed to farm radio programming early on would have experienced a growth of approximately 8% on farm land value per acre over after two decades relative to a similar county without farm radio by the end of this same period. The effects on crop value per acre are also positive on the short run but dissipate by the end of the event window.

During the sample period, the quality of radio receivers available in households changed dramatically due to new technologies and to changes in the demand for portable radio sets as rural electrification programs brought electric power to farm families. These changes over time could have an unpredictable effect on the dichotomous treatment assignment based on a threshold of signal strength. Because of this limitation, together with the pre-trends of panel (b) in Figure 1.4, I interpret with caution these results as suggestive evidence that farm radio programming had lasting effects on agricultural productivity.

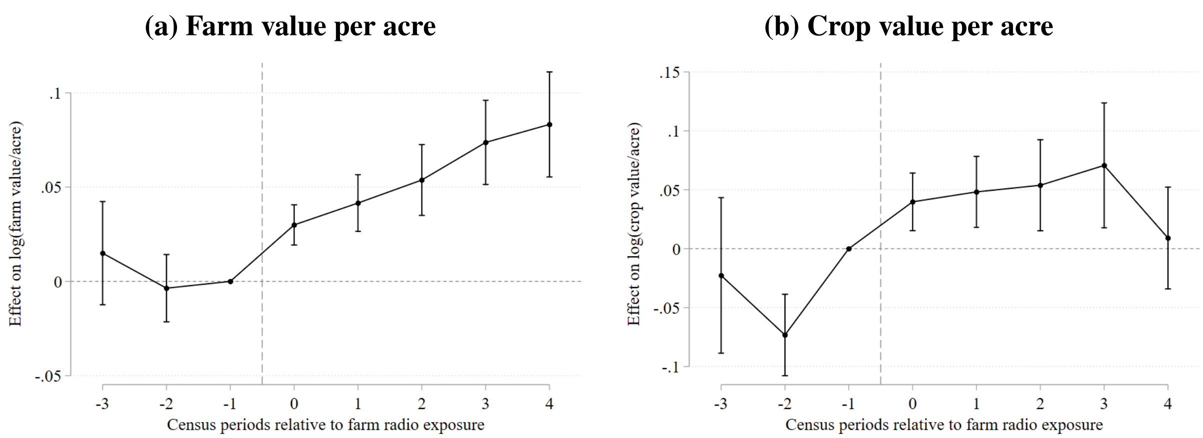



Figure 1.4: Cumulative effect of farm radio exposure

*Notes:* The figures show point estimates and 95% confidence intervals of the $\beta_\ell$ parameters from Equation 1.2, representing the dynamic cumulative effect of $FarmSignal$ on the outcomes of log of farm value per acre and log of crop value per acre on panels (a) and (b) respectively. The period prior to farm radio exposure is normalized to zero. Standard errors are corrected for clustering at the county level.

**Economic significance and comparison with related literature.** American agriculture underwent significant transformation during the time period here studied, and in the 1930s, in certain regions of the country, farmers faced adversity due to drought and economic pressures of the Great Depression. During the economic recovery period from 1940 to 1950, counties in the main analysis sample experienced real growth of 23.8% and 45.3% in the per acre value of farm land and of all crops respectively. Put into perspective, assuming the short-run effects of farm radio are homogeneous across counties, my estimates of 2.1% for farm value and 4.4% for crop value imply a counterfactual increase in signal strength (i.e. radio availability) of one standard deviation for all counties would amount to 8.8% and 9.7% of the 1940 to 1950 growth on farm value and crop value per acre respectively.[11]

Prior studies look at other important drivers of agricultural land prices and productivity, such as credit constraints, market access, and information barriers. Hutchins (2023) examines an expansion of the Farm Credit System in the U.S. and find that counties that benefited from improved improved credit access experienced a 7% to 14% increase in crop value per acre. Chan (2022) examines the expansion of American railroads and find that a one standard deviation increase in market access increased agricultural land values by .23 standard deviations (or 31%, using the paper's summary statistics). Closest to my research design but using a more modern context, Gupta et al. (2020) study the expansion of mobile phone networks in rural India, coupled with call centers providing agricultural advice. They find that areas with a one standard deviation increase in mobile phone coverage and no language barriers between farmers and agricultural advisers experienced a 1.3% increase in yields for a composite of major crops in the region. In relation to these studies, my findings suggest that farm radio had an important (albeit arguably not the biggest) role in transforming agriculture in the 20$^{th}$ century.

## 1.5 Channels

I now explore possible channels that might explain the short-run effects of farm radio documented on the previous section.

**Other Radio Stations.** A possible explanation for the main result presented on farm value is that the agricultural land prices reflect more than just productive value. A potential concern is that the main results could reflect exposure to radio programs in general. Radio offers consumers an amenity through general programming unrelated to information provision to farmers, and such benefits may be reflected on land prices. I explore this channel utilizing data on "other" radio stations that place less emphasis on locally targeted farm content, i.e., stations not included in the curated list from the State Agricultural Radio Programs in Brunner (1936).

[11] The caveat being that in reality the effects of increasing signal strength are most likely diminishing. This counterfactual increase should be thought of as going from little to high radio availability over all counties.

Table 1.4 reports the results from estimating a statistical "horse race" version of equation 1.1 after adding to the richest specification the other radio stations' ITM-predicted and free space-predicted signal strengths, represented by the $OtherSignal$ and $OtherSignalFree$ variables in the table. While the results on farm radio remain unchanged, I find precisely estimated null effects on overall agricultural productivity, measured both by farm and crop value per acre. These results strengthen the interpretation of the previous subsection that targeted farm radio programming specifically drove productivity growth in the agricultural sector.

Table 1.4: Robustness Check – Exposure to Other Radio Stations

| | (1)<br>Log(farm value/acre) | (2)<br>Log(crop value/acre) |
|---|---|---|
| $FarmSignal$ | 0.022*** | 0.044*** |
| | (0.005) | (0.009) |
| $OtherSignal$ | 0.001 | -0.008 |
| | (0.005) | (0.010) |
| County fixed effects | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ |
| $FarmSignalFree$ | ✓ | ✓ |
| $OtherSignalFree$ | ✓ | ✓ |
| Baseline controls | ✓ | ✓ |
| Soil characteristics × year controls | ✓ | ✓ |
| Observations | 13,380 | 13,380 |
| Number of Clusters | 2,230 | 2,230 |
| Adjusted R-Squared | 0.958 | 0.829 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* $FarmSignal$ is a standardized measure (mean zero and variance one) of the predicted signal strength of farm radio resulting from the Irregular Terrain Model. $FarmSignalFree$ in contrast is the predicted signal strength assuming a smooth and featureless earth. $OtherSignal$ and $OtherSignalFree$ are similarly defined signal strengths for other (non-farm targeting) radio stations. Baseline controls include log of total population and farm population, percentage of farms with tenancy regime, and percentage of males, Black individuals, and manufacturing workers. Soil characteristics are time-invariant, and include soil water capacity, % of soil consisting of clay, soil erodibility (K) factor, soil drainage quality, liquid limit of the soil layer, and soil annual flood frequency. Farm value per acre is the ratio of the combined value of all farms and buildings over the acres of farm land and crop value per acre is the ratio of the aggregate value of all crops over the acres of harvested cropland. Standard errors are corrected for clustering at the county level.

Next, I explore possible differential effects of farm radio due to county characteristics that may influence agricultural productivity. To do so, I add to equation 1.1 an interaction of the ITM-predicted signal strength of farm radio with a variable of interest. For ease of interpretation, these added variables discussed below are also standardized such that they have a mean of zero and standard deviation of one.

**Information barriers.** Farm radio may have larger benefits for farmers facing higher costs for information acquisition. I test this hypothesis by examining the interaction between farm radio signal and the signal of other radio stations, here acting as a proxy for media saturation. Columns (1) and (5) of Table 1.5 show that this interactive effect is negative, though only statistically significant for the farm value per acre measure, suggesting that farm radio was particularly helpful for farmers with less access to alternative sources of information. Farmers may also benefit more on areas with less knowledge flows. I utilize data on railroad networks in 1911 from Atack et al. (2010) to compute the distance from county centroids to the nearest segment of railroad, which I then interact with farm radio signal. Columns (2) and (6) show a strong and positive interaction effect, suggesting that more isolated areas with less transportation infrastructure received larger gains from farm radio.

**Human capital and economic status.** I now explore differential effects derived from demographic characteristics in 1930. Literature dating back to Nelson and Phelps (1966) posit that education can remove barriers to knowledge diffusion.[12] I examine the interaction between farm radio and illiteracy rates in 1930 and find on columns (3) and (8) of Table 1.5 that the effects of farm radio were larger among the less educated, although the effect is small and insignificant for farm value per acre. I similarly examine farm radio's interaction with economic status, proxied by occupational income score.[13] On one hand, farmers of lower economic status may have higher marginal returns for technology adoption. On the other hand, farmers with higher economic status face lower liquidity constraints and are able to make productivity-enhancing capital investments. Columns (4) and (9) show inconclusive results where the per acre effect of farm radio on farm

[12]More recent work by Squicciarini and Voigtländer (2015) explore this idea on the context of upper-tail education.

[13]Direct measures on educational attainment and income are not available until the 1940 census. I use the share of illiterate among the population aged ten and above as a proxy for education level. I use county averages of the 1930 occupational income scores (sourced from the 1930 census microdata available at IPUMS), which is commonly used in studies of labor market outcomes from this era (Saavedra and Twinam, 2020).

value is significantly larger in areas with a lower occupational income score, but insignificant and of opposite sign for crop value.

Table 1.5: Potential Channels

| | Log(farm value/acre) | | | | Log(crop value/acre) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $FarmSignal$ | 0.018*** | 0.021*** | 0.020*** | 0.019*** | 0.042*** | 0.044*** | 0.037*** | 0.044*** |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.009) | (0.009) | (0.009) | (0.009) |
| $FarmSignal \times OtherSignal$ | -0.006*** | | | | -0.004 | | | |
| | (0.002) | | | | (0.003) | | | |
| $FarmSignal \times RailroadDist$ | | 0.016*** | | | | 0.014*** | | |
| | | (0.003) | | | | (0.004) | | |
| $FarmSignal \times \%Illiterate$ | | | 0.004 | | | | 0.026*** | |
| | | | (0.003) | | | | (0.005) | |
| $FarmSignal \times OccScore$ | | | | -0.011*** | | | | 0.003 |
| | | | | (0.002) | | | | (0.004) |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| County fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $FarmSignalFree$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Baseline controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Soil characteristics × year controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 13,380 | 13,380 | 13,380 | 13,368 | 13,380 | 13,380 | 13,380 | 13,368 |
| Number of Clusters | 2,230 | 2,230 | 2,230 | 2,228 | 2,230 | 2,230 | 2,230 | 2,228 |
| Adjusted R-Squared | 0.958 | 0.958 | 0.958 | 0.958 | 0.829 | 0.829 | 0.829 | 0.829 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes:* All baseline variables defined as previously in Tables 1.2 to 1.4. $RailroadDist$ measures the distance from county centroids to the nearest segment of the railroad network in 1911. $\%Illirate$ is the percentage of illiterate population aged ten and above in 1930. $OccScore$ is the occupational income score in 1930 from individual census microdata in IPUMS, averaged at the county level. These three interacted variables are standardized with a mean zero and variance one. Standard errors are corrected for clustering at the county level.

## 1.6 Additional Results

**Alternative samples.** The residual variation in signal strength from the ITM versus free space model can become larger as the distance increases between radio stations and county centroids. As such, this residual variation is minimized when the station is located inside the county, in which case controlling for free space signal may not fully address the concern of endogenous station location.

Geography also matters for the suitability of different crops and the characteristics of farms. To examine the importance of these threats to identification, I re-estimate the impact of farm radio on agricultural productivity across different samples.

Figure 1.5 reproduces the estimates for farm value per acre on panel (a) and crop value per acre on panel (b) on various samples of counties. The baseline estimates on the top of each panel replicate column (4) of Table 1.2 and column (1) of Table 1.3 respectively. Moving downwards, I show the estimated coefficient of $FarmSignal$ from equation 1.1 on samples comprising counties with a high suitability for growing a specified crop, where a county-level crop suitability index is constructed from the FAO gridded data. Counties are considered highly suitable if the index is above 50%, a threshold which McGowan and Vasilakis (2019) find correlates positively – in the context of corn – with the probability that the crop is grown in the county. In my estimates, we see coefficients that are somewhat stable and comparable, if not larger in magnitude, with the full baseline sample, with the exception of cotton where the coefficients are in addition estimated with less precision due to the smaller sample size as this crop is predominantly grown on the American South. While the results are generally consistent with the baseline, it is worth noting that they are highest in magnitude on counties suitable for growing wheat. This could be partly due to the salience of wheat on farm radio programming, which is illustrated by a word cloud in Appendix Figure A3 of scientific terms constructed from transcripts of the *National Farm and Home Hour*.[14] Another possible explanation for the slightly larger estimates of aggregate productivity in counties with high suitability for the grain crops more generally is that information from farm radio can lead to better decisions on which crops to grow and where, improving resource allocation.[15]

Moving further down on the figure, we see that estimates are almost identical to the baseline for the sample of counties more than 100km away from farm radio stations in any period of the data, where farm radio exposure is even more likely to be exogenous. At the bottom of the figure we see that the main results are also robust to including all counties in the continental US, ignoring the

[14]As it refers to a nationwide program, the content of the *National Farm and Home Hour* may not accurately reflect farm programming at a local level.

[15]A strand of literature inspired by the work of Hsieh and Klenow (2009) relates total factor productivity in agriculture to misallocation of resources. See Le (2020) and Chen et al. (2023) for two recent examples.

issues of outliers introduced by measurement error and highly urbanized counties. The different samples shown in this figure can be visualized in maps shown in Appendix Figure A4.

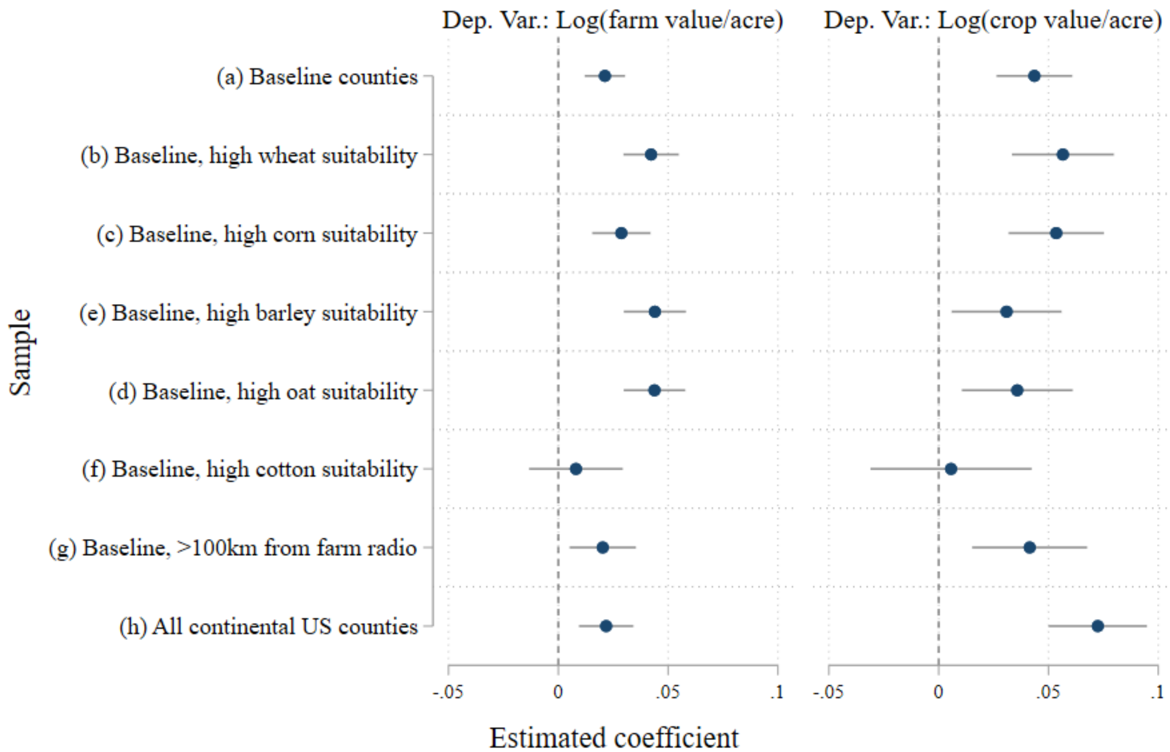



Figure 1.5: Alternative Samples

*Notes:* Plotted estimated coefficients are for $FarmSignal_{ct}$ in the full model outlined by equation 1.1. The grey lines represent 95% confidence intervals. See Section 3.1 for sample selection leading to the baseline counties in estimate (a). High crop suitability in estimates (b) to (f) is assigned to counties with a crop suitability index above 50% using county averaged data from the FAO gridded suitability index. Estimate (g) drops counties within 100km of any farm radio station in any sample year. Estimate (h) includes all counties in the continental US.

**Change in remoteness.** While radio may help integrate farmers in remote communities, an alternative explanation for my main results is that other investments in infrastructure may co-occur with the increasing availability of radio. I use data from Burghardt et al. (2022) on the evolution of road networks to assess whether improvements in road infrastructure over this relevant time period. Table A2 shows the result of adding to equation 1.1 an interaction of ITM-predicted signal strength of farm radio with $NewRoadsDensity$, the number of new roads per square kilometer of built area in counties from 1900 to 1940, standardized with a mean of zero and standard deviation of one. The estimated effect of farm radio changes slightly as the sample of counties is here restricted due to limited availability of the road infrastructure data but remain qualitatively unchanged. The coefficients on the interaction variable are statistically insignificant and economically small, suggesting radio remained a relevant driver of agricultural change even for counties becoming better connected by road networks.

**Sensitivity to antenna height missing values.** The identification comes from variation in signal strength due to topography which is measured by the irregular terrain model. Since antenna height – a key parameter on the signal strength prediction – is only available in one year of data, I here assess the sensitivity of the results to alternative approaches to handling missing antenna height values in the data. To do so, I recalculate the strongest farm radio signal strength[16] of each county in two alternative ways: (1) replacing the missing antenna height values with the median antenna height of 250ft from the available data in the 1940 issue of the *Broadcasting Yearbook*, and (2) replacing all antenna height values in all years with this same value of 250ft. In the data, the correlation between the baseline signal strength and the recalculated signal strength under different antenna height assumptions is over 0.99. Unsurprisingly, Appendix Figure A2 shows that the main results are virtually unchanged to different strategies to address the problem of missing antenna height values.

**Binary signal strength.** In alignment with the event study evidence presented in section 4, I perform an additional robustness check in Table 1.6 where the continuous measure of signal strength $FarmSignal$ gets replaced by an indicator equaling one if the ITM-predicted farm radio signal is at or above the 1925 median and zero otherwise. While the estimates remain qualitatively similar, the larger estimated coefficients obtained with the binary measure suggest the effect of signal strength is unlikely to be linear. Without information on the technical characteristics of farmers' radio receivers, it is difficult to pin down precisely the threshold of usable signal strength and improve upon $FarmRadio$ as a proxy for farm radio exposure. This limitation also highlights the fact that the residual variation in signal strength in my model is being used to identify the intent-to-treat

[16] Only the signal strength calculated with the irregular terrain model is affected to changes in this parameter. The antenna height input has no effect on the free space signal propagation since this model assumes there are no topographical features in the line of sight between the transmitter and receiver.

effect of mass media (Crabtree and Kern, 2018).[17]

Table 1.6: Robustness check: Binary Signal Strength

| | (1) Log(farm value/acre) | (2) Log(crop value/acre) | (3) Log(wheat yield) | (4) Log(corn yield) | (5) Log(barley yield) | (6) Log(oat yield) | (7) Log(cotton yield) |
|---|---|---|---|---|---|---|---|
| $\mathbb{1}[FarmSignal > \mu_{1/2}^{1925}]$ | 0.024*** | 0.057*** | 0.100*** | 0.012 | 0.113*** | 0.088*** | 0.161*** |
| | (0.008) | (0.014) | (0.015) | (0.012) | (0.020) | (0.013) | (0.019) |
| County fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $FarmSignalFree$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Baseline controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Soil characteristics × year controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 13,380 | 13,380 | 11,859 | 13,257 | 9,872 | 12,907 | 5,109 |
| Number of Clusters | 2,230 | 2,230 | 2,091 | 2,226 | 1,894 | 2,218 | 878 |
| Adjusted R-Squared | 0.958 | 0.829 | 0.616 | 0.792 | 0.486 | 0.598 | 0.685 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes:* $\mathbb{1}[FarmSignal > \mu_{1/2}^{1925}]$ is an indicator equaling one if the predicted farm radio signal strength meets or exceeds $\mu_{1/2}^{1925}$, the 1925 median. All other variables defined as previously in Tables 1.2 to 1.4. Standard errors are corrected for clustering at the county level.

**Alternative specifications.** Lastly, I perform additional sensitivity checks in Appendix Table A3 showing in columns (2) and (5) that the main results are robust to flexibly controlling for free space signal propagation with a cubic polynomial that allows for a nonlinear effect of proximity to radio station. Columns(3) and (6) show the results are also quantitatively similar after weighting the regression with the county's farm population.

## 1.7 Conclusion

I provide evidence that early radio stations that worked in collaboration with universities and agricultural experiment stations to broadcast farm programming had a measurable and persistent impact on agricultural outcomes. These impacts were more pronounced among disadvantaged farmers residing in counties with lower literacy and economic status and lower access to markets. The effects were felt across many of the most prominent crops grown in the country, and also captured by overall productivity measures related to land prices and total crop revenues. Still, a

[17] Importantly, this proxy remains policy relevant as the availability of radio stations can be manipulated through investments in broadcasting infrastructure.

limitation of this study is the lack of key variable inputs, such as seed varieties, which would allow us to explore the importance of farm radio on the adoption of productivity-enhancing technologies.

The findings in this paper are relevant to the policymakers of today, who are searching for cost-effective alternatives to remove information barriers to farmers in developing countries. Despite being a century-old technology, radio remains an affordable, long-reaching, easy-to-use, and relevant source of information (over 55% of sub-Saharan African households still tune in weekly, according to Aker (2011)). Previous findings of limited effectiveness of information offered by radio could be attributed to the lack of commercial incentives to provide locally targeted content to farmers. This issue can be overcome with government-sponsored programming, where radio broadcasters partner with extension services and research institutions to deliver locally relevant information in regions lagging in agricultural productivity.

# Chapter 2

# Clouded Thoughts: Air Pollution and Cognitive Performance

## 2.1 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. 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 Ebenstein et al. (2016), 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 high-profile players such as Lee Sedol from Korea and more recently Ke Jie from China were defeated by Google's AIs Alpha Go and Alpha Go Zero respectively.[1] I analyze records of high-level Go games using Leela Zero – an open-source AI modeled after Alpha Go Zero – as an "expert evaluator" which classify players moves as strong, acceptable, or inaccurate. At the time of writing this paper, Leela Zero has already defeated 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 game 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, providing decades of annotated game data. Tournaments are typically played indoors in hotels and game salons, in controlled environments, mitigating 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

[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 professional player Hajin "Haylee" Lee in 8 out of 8 games. The games are available at *Haylee's World of Go/Baduk."*

[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).

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. As players increase in rating (i.e., become stronger), their percentage of strong moves per game on average increases and percentage of inaccuracies decreases. Moreover, in a match-up of similarly skilled players, the winning odds are 1.52 for the player who makes more strong moves in a game and similarly the winning odds are 0.35 for the player who makes more inaccuracies.[4] 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 35% 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.[5] Other pollutants, namely $O_3$, $SO_2$ and $CO$, see only a modest increase and in some cases a small decrease in concentration during dust days, which suggests that particulate matter pollution is a key driver of the main results below.[6]

Finally, using a fixed effects model I estimate the effect of Asian dust on the cognitive performance outcomes. The main finding suggests that exposure to Asian dust increases a player's propensity to make inaccurate choices. Players overall make 8.3% more inaccurate moves when exposed to Asian

[4]In plain English, this amounts to the player making more strong moves winning 3 out of 5 games and similarly the player making more inaccuracies winning just 7 out of 27 games.

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

[6]A variety of toxic materials are found in Asian dust and I cannot rule out the possibility that a different pollutant drives the results. However, the Japanese data contains a few additional pollutants for which the concentrations are not significantly affected by dust storms.

dust during the game day, which amounts to roughly one additional inaccuracy per one hundred decisions on the move to be played.

For a comparison on a fairly different cognitive skill, 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 $PM_{2.5}$. While my estimated effect appears larger, the increase in particulate matter exposure (both coarse and fine) induced by Asian dust is also significantly larger. Contrary to my results, recent work by Holub and Thies (2023) find no significant effect on the frequency of severe errors on the coding output of knowledge workers on GitHub when exposed to higher $PM_{2.5}$ pollution levels. The authors explain that code developers work on flexible settings that allow for adaptation, and thus they react to days of high pollution concentration by switching to easier tasks and by working fewer hours.

I uncover some heterogeneity after reproducing these estimates for the subpopulations below and above the sample median of 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 inaccuracies during the induced air pollution shock or nearly two additional inaccuracies per one hundred moves. These heterogeneous effects are consistent with a strand of health literature highlighting the susceptibility of older individuals to adverse effects of air pollution. For instance, Lai et al. (2022) estimate that large shocks in particulate matter pollution from straw burning impairs the cognitive functioning of individuals aged 55 or older in dimensions such as mental intactness, immediate recall, and delayed recall. Krebs and Luechinger (2021) similarly find suggestive evidence that exogenous exposure to nitrogen oxides reduces performance in a brain training game, and the adverse effects are stronger on players aged 60 or older.

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 broader population – play nearly two additional inaccuracies per one hundred moves.

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). A plausible 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 demands a high degree of inductive reasoning and is performed in a controlled 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.[7]

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

[7]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.

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.

## 2.2 Literature Review

### Air Quality, Health, and Labour Productivity

There exists a vast literature that study the relationship between the environment and the population well-being.[8] 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. 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 has 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 worker productivity by 5.5% for 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 implication for workers engaged in tasks that are mostly, or purely, cognitive.

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

Two recent papers investigating the impact of pollution exposure on different mental faculties are closely related to my research. First, Ebenstein et al. (2016) 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 broad measure of cognitive ability (university test scores) for an economically relevant population of students. My research complements Ebenstein et al. (2016) 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 inaccuracies in strategy board games, but they nonetheless provide useful estimates against which I can benchmark my results.

## 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, scientific research such as Lee et al. (2007) has found in this dust 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. 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 finds suggestive evidence of a behavioural response where mortality effects decrease due to dust advisory.

Jia and Ku (2019) 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.

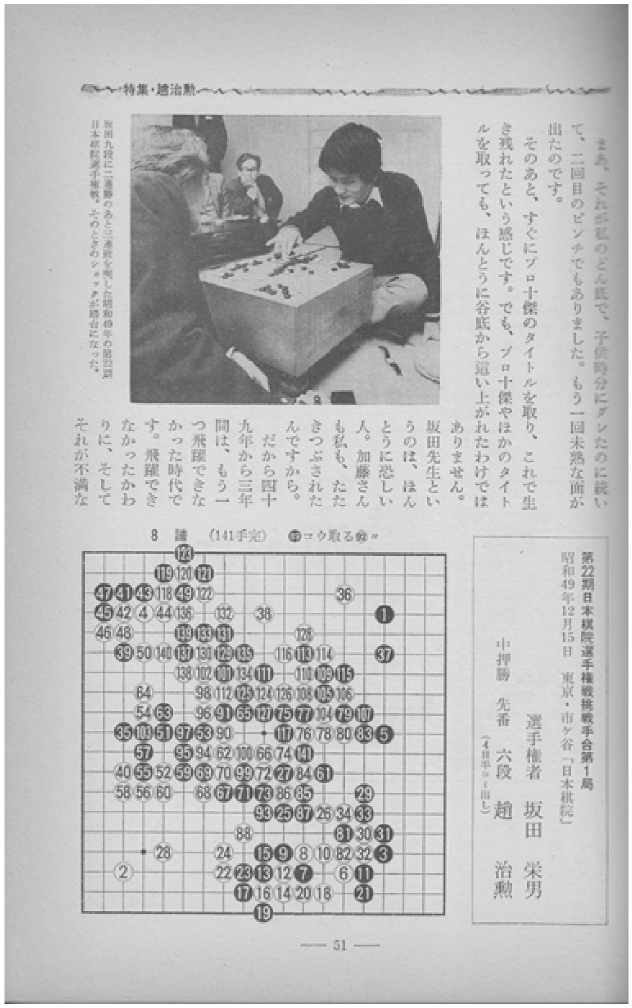
While there is strong evidence supporting the health-related effects of the Asian dust, no attention has been given to the possibility that this phenomenon may affect economic outcomes through short-term deterioration of cognitive functioning. PM pollution can penetrate into the lungs and, if the particulate 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 may have important policy implications.

## 2.3 Go

### A Primer on the Game

The history of Go dates back at least to 300 B.C., with Chinese scholars such as Confucius utilizing the game to illustrate thoughts about human nature. The origins of Go as a competitive board game can be traced back to the formation of the Nihon Ki-in (Japan Go Association) in the 1920s. The Japanese professional system was brought to Korea in 1945 with the formation of the Hanguk Kiwon (Korea Baduk Association) by Cho Namchul, who studied the game in Japan. The establishment of the Japanese and Korean associations led to many newspaper-sponsored tournaments, with interesting sequence of moves or even entire games being published and analyzed by the major media outlets within these countries. The associations also issue rank certificates that serve as a proxy of player ability: a player is deemed a professional after reaching the 1-dan level. 9-dan is the highest ranking that can be achieved. Similarly, amateurs are ranked from 1-amateur dan to

9-amateur dan. Informally, professionals ranked between 5-9 dan are regarded "high dan," while 1-4 dan professionals are considered "low dan." Figure 2.1 shows a magazine excerpt which reproduces a game record between two high dan players at the Nihon Ki-in; the numbering on the diagrams corresponds to the order in which the moves were played.



特集・趙治勲

まあ、それが私のどん底で、子供時分にグレたのに続いて、二回目のピンチでもありました。もう一回未熟な面が出たのです。

そのあと、すぐにプロ十傑のタイトルを取り、これで生き残れたという感じです。でも、プロ十傑やほかのタイトルを取っても、ほんとうに谷底から這い上がれたわけではありません。

坂田先生というのは、ほんとうに恐しい人。加藤さんも私も、たたきつぶされたんですから。

だから四十九年から三年間は、もう一つ飛躍できなかった時代です。飛躍できなかったかわりに、そしてそれが不満な

坂田九段に二連勝のあと三連敗を喫した昭和49年の第22回日本棋院選手権戦。そのときのショックが踏台になった。

8 譜 (141手完) ⓭コウ取る⓼〃

第22期日本棋院選手権戦挑戦手合第1局
昭和49年12月15日 東京・市ケ谷「日本棋院」

選手権者 坂田 栄男
中押勝 先番 六段 趙 治勲
(4目半コミ出し)

— 51 —

Figure 2.1: Title match game of the 22nd Annual Nihon Ki-in Championship.

The game is played on a 19×19 grid with the following basic rules. There are two players, Black and White, who take turns either placing a single stone with their respective colours on the vacant board intersections or passing a move. The first move is always given to the Black player. 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. The game concludes when both players pass and the winner is determined by counting one point for each territory and captured stone of each player's and adding *komi*, compensation points given to White for Black's first mover advantage [9].

Go players often use jargon to describe common board patterns and game situations. The opening stage of the game is known as *fuseki*, typically involves moves near the board four corners, and can last up to fifty moves according to most opening theory sources. Unlike chess, openings in Go are not very systematic since early moves are commonly made in isolation, with little contact with the opponent's stones. Systematic play occurs more frequently as the game progresses, where stone placements by both players offer plenty of opportunities for territorial disputes yet the board still offers a large number of plausible move choices. Many move sequences have been documented in high-level games; when sequential moves are considered strong and balanced for both sides they are characterized as *joseki* and become an object of study of many Go players. High level players display an impressive knowledge of *joseki* and often learn about potential weaknesses in opponents' strategic patterns by studying their game records.

An important feature for this analysis is that sequence variations within a *joseki* are known to lead to different positional advantages for each player, which means careful decision-making during these sequences are key for achieving victory in the game. Playing Go at a high level is not a trivial task: the action space is very large at any node of the game, since practically all open spaces constitute legal moves. Stone placement can have an obvious impact on a local territorial dispute but also inconspicuous global effects on the game board. Even professional players cannot always articulate what characterizes a good move, often resorting to ambiguous terms for describing moves qualitatively. For example, Go players denote by *tesuji* a "strong" play – the best move in a local position – and by *poka* an inaccurate play, both of which are terms reflecting an imperfect decision-making process due to cognitive limitations of our minds as well as the time constraint for making decisions. Such bounded rationality problem implies that Go players are unable to evaluate, with enough precision, the outcomes associated with every possible choice and thus make decisions based on adequacy rather than the true optimal solution. This "satisficing" heuristic suggests that

[9] 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.

expert players are forward-looking and conduct, at each node of the game, a highly selective tree search for the optimal move (Igami, 2020).

## Computer Go

Go has been a subject of computer science research for decades. However, recent advances in machine learning and Monte Carlo tree search (MCTS) methods in the 21st century led to remarkable progress in the strength of AI in the game, which went from amateur level in 2014 to defeating the reigning world champion Ke Jie in 2017. A major breakthrough in computer Go was Google DeepMind's Alpha Go AI introduced in 2015 (Silver et al., 2017). Alpha Go combines a deep neural network (DNN), trained with human expert moves from high-level game records, with a tree search algorithm that evaluates the strength of candidate moves recommended by the DNN. In 2017, DeepMind released Alpha Go Zero, a new and much stronger Go-playing bot based on an unsupervised learning algorithm. The unsupervised nature of the learning means this AI did not receive knowledge from human games, unlike its predecessor Alpha Go.

Alpha Go Zero is trained from self-play games by a reinforcement learning algorithm. Its logic is as follows. At each node of a game, the AI's latest DNN evaluates the set of legal moves using a policy function and formulates a set of candidate moves from choices achieving the highest Policy Network (PN) scores.[10] Next, the MCTS algorithm performs a large number of simulations starting from each candidate move, where each simulation traverses the game by making moves following a stochastic process distributed with an exogenously given prior. At the terminal node of each simulation, a game winner and the game history are stored. After all simulations, each candidate move is assigned a Value Network (VN) score based on the number of "Monte Carlo wins" that follow each proposed move and the move with the highest VN score is chosen by Alpha Go Zero. The number of visits each game node receives during the simulations is also stored as the MCTS "search probability." Before any training, the neural network is initialized with random weight parameters. After each self-play iteration, these weight parameters are updated to approximate the

[10]The policy network output can be thought of as the highest-ranked predicted conditional choice probabilities using the first step estimation of Hotz and Miller (1993).

move probability and value of the neural network with the MCTS search probabilities and self-play VN score.

Prior to its deployment against human players, Alpha Go Zero trained without human intervention for 4.9 million games using 1600 simulations for each candidate move at the MCTS step. Over the course of its learning, the AI discovered many fundamental aspects of human knowledge of Go, although it also developed unconventional strategies beyond our current understanding of the game. Importantly, the AI learned many professional *joseki* patterns and grasped sophisticated Go concepts such as stone grouping, territorial influence, and *sente* (forcing moves).

In this paper, I rely on the expert opinion of the Leela Zero AI which is an open-source engine modeled after Alpha Go Zero. Just like Google DeepMind's AI, Leela Zero's style of play aligns with human knowledge of the game. Four facts highlight Leela Zero's usefulness for my analysis: its unsupervised learning nature mitigates the risk of a gameplay contemporaneous to a training dataset, as may be the case for AIs learning through human games; second, the use of tree search methods over a constrained set of candidate moves resembles the behaviour of human experts; third, its ability to perform thousands of computations per second makes it feasible to bulk analyze a large number of games in a reasonable time; fourth, the move assessments contemplate only maximizing the odds of an eventual win and involve no bluffing or other risky playing strategy. In light of these facts, I use Leela Zero's move recommendations to estimate the percentage of strong moves and inaccuracies for each game and player and claim that these are relevant *objective* measurements of quality of decision-making. Before elaborating on the empirical strategy, the next section provides an overview of the air quality data and Go game records used in this paper.

## 2.4 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 to air quality and game records are separately discussed below.

## 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 it is known to the Japanese) is available daily from 59 weather stations at the Japan Meteorological Agency's (JMA) website starting from 1967. Figure 2.2 presents a map of the 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).

## Game Data

A number of books dedicated to the study of Go are published every year, creating a large paper-based archive of historic games. For this research, historical records of Go were sourced from the Games of Go on Disk (GoGoD) database. GoGoD is an online collection of games sourced from printed

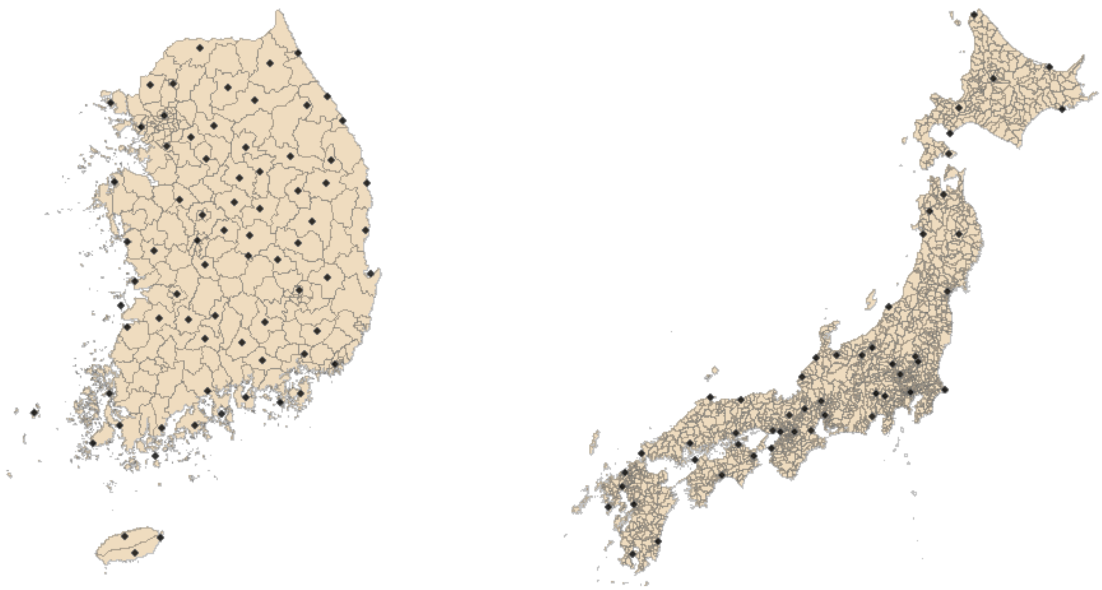


Figure 2.2: Asian Dust Stations in South Korea (left) and Japan (right)

and online media, currently containing 96,800 games played between the years 196 A.D. and 2018. 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 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. Because dan degrees represent an achievement, they may inaccurately represent player ability years after the degree award. I complement the data set with Bayesian Elo ratings obtained from Go Ratings (Coulom, 2020) that capture player strengths at the day a game is played.[11] I also complement the information available in the SGF files with individual characteristics sourced from player biographies that accompany the GoGoD database. Table 2.1 displays the combination of properties in SGF files and sourced player characteristics that are used in the analysis section.

The dataset used in the analysis consists of 22,213 games for which all variables listed in Table 2.1 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

[11] See Coulom (2007) for details on the theory and estimation of Bayesian Elo ratings.

Table 2.1: Variables Extracted from Game Records, Player Biographies, and Online Sources

| **Property name** | **Description** |
|---|---|
| Black Name | Name of Black player |
| White Name | Name of White player |
| Black Rank | Rank of Black player |
| White Rank | Rank of White player |
| Black Elo | Elo of Black player |
| White Elo | Elo of White player |
| Black Age | Age of Black player |
| White Age | Age of White player |
| Black Gender | Gender of Black player |
| White Gender | Gender of White player |
| Moves | Number of moves played in game |
| Date | Date of game |
| Place | Place where game was played |
| Event Name | Name of game event |

age at game date distribution ranges 11 to 97 years old players, with median and mean at 30 and 33 years, respectively. 9-dan (the highest degree achievable) players amount to 47% of the individuals.

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 2.2, 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.

[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.

Table 2.2: 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) | |

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

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.

The AI's prior knowledge from training data combined with the Monte Carlo simulations result in a set of candidate moves, each of which receives a VN score between zero and one, where the

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

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 preferred computer move; and, (2) the percentage of moves played that are not in the set of candidate moves proposed by Leela Zero. Leela Zero's playing strength currently exceeds the strongest Go professionals who now use AI as a training and game review tool. In the game's idiom, players can use AI to identify a move as *tesuji* (strong move) and *poka* (blunder move), both of which are related to the constructed measures of strong and inaccurate moves. In the analysis, the outcome variables constructed from (1) and (2) are denoted $\%strong$ and $\%inaccurate$ respectively.

## 2.5 Empirical Strategy

I exploit the random and short-lived nature of the Asian dust storms to investigate the relationship between air pollution and the cognitive performance of Go players. This is done by examining how temporal and spatial variation in air quality induced by Asian dust storms in South Korea and Japan affect the quality of decision-making of Go professionals and skilled amateurs exposed to this exogenous air pollution shock.

My analysis focuses on the short-term impact of air pollution on cognitive performance, abstracting from any long-lasting effects of cumulative exposure. This means incidence of Asian dust will be thought of as capturing only a short-term effect on decision-making of the variation in air pollution during the day of a dust storm. With this in mind, the analysis is kept at the game level where we can exploit the daily regional variation of dust incidence. Quality of decision-making is modeled through the outcome variables pertaining move strength $\%strong$ and $\%inaccurate$ defined in the previous section. The main specification of the analysis estimates the effect of Asian dust events on the outcome variables after controlling for potential confounders:

$$Y_{pjt} = \alpha + \delta Dust_{jt} + \beta X_{pt} + \eta_{y(t)} + \eta_{m(t)} + \psi_j + \phi_p + \varepsilon_{pjt} \tag{2.1}$$

[14]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. Appendix Figure B1 plots the histogram of candidate moves per node in my parsed output of move evaluations by Leela Zero.

where $Y_{pjt}$ refers to either the percentage of strong moves or inaccurate moves played by individual $p$ on a game that took place on city $j$ and day $t$. $Dust_{jt}$ indicates a dust event on city $j$ and day $t$; $X_{pt}$ is a vector of controls which may be included in some specifications. These controls are $Female_p$ which equals one if the player is a female and $Strength_{pt}$, which is the estimated Elo rating of the player on game day. $\eta_{y(t)}, \eta_{m(t)}$, $\psi_j$, and $\phi_p$ are year, month, city, and player fixed effects which may be included on certain specifications.

The coefficient of interest $\delta$ uncovers the effect of an Asian dust day on the quality of decision-making. Player sex accounts for any possible differential responses of males and females on the probability of playing a game in a dust day (one can think of a differential health effect between males and females). Elo rating is a proxy for player strength, and it controls for possible pollution avoidance behaviour leading, for example, to rescheduling of games in low-profile tournaments that cater to lower-ranked players. Year and month fixed effects control for time trends in air quality to make the pollution shocks induced by Asian dust comparable over the sample period.[15] City fixed effects account for geographical and geological features that correlate both with Asian dust incidence and the sorting of players into different cities. The richest specification includes player fixed effects, exploring within-player variation in Asian dust exposure. Errors are clustered at the city level, which is the geographic unit of exposure for Asian dust events.

I also regress the main specification separately for the subpopulations below and above the median age in the sample (see Figure 2.3 below for the age distribution of players in the sample). This is done in light of evidence of heterogeneous responses to air quality in the health effects environmental literature. The coefficient of interest $\delta$ represents the causal effect of Asian dust on the probability a player makes a strong move or inaccuracy under the assumption that unobserved factors jointly affect the outcome measurements of cognitive performance and the probability of a player getting exposed to an Asian dust event. This assumption could be violated, for example, if players with a poor health condition that hinders mental acuity are more likely to not show up for a scheduled tournament during a dust day.

[15] Lee et al. (2013) shows air pollution levels in major cities in South Korea have decreased over the course of time.

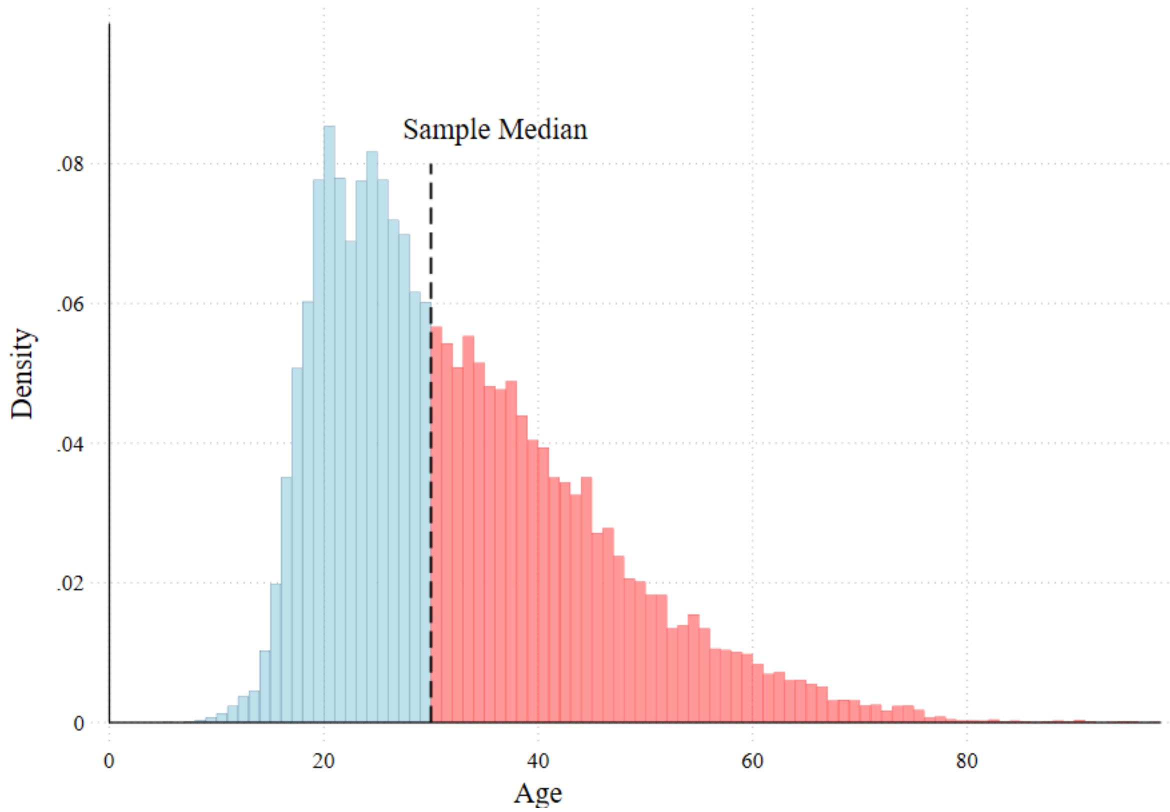



Figure 2.3: Age distribution of players in the sample

*Notes:* Sample comprises the 22,213 games for which all variables listed in table 2.1 are available. The age of each player is inferred from the date of birth found in the player bibliographies accompanying the GoGoD database.

## 2.6 Results

I first examine the association between the outcome variables and player strength. Figure 2.4 presents a scatter plot the mean of $\%strong$ across each Elo percentile of player strength, ordered from weakest to strongest. The graph shows that the strongest players are also the most likely to play Leela Zero's best suggested move, with the percentage of strong moves per game increasing over 5 percentage points from weakest to strongest players.

Similarly, Figure 2.5 presents the relationship between the mean of $\%inaccurate$ across Elo percentiles, and the same intuition from Figure 2.4 holds: players become less likely to make inaccuracies as their estimated playing strength increases. The strongest players make 4p.p. fewer inaccuracies than the weakest players in the sample.[16]

[16] Appendix Figures B2 and B3 show an analogous version of Figures 2.4 and 2.5 using the dan degree achieved by

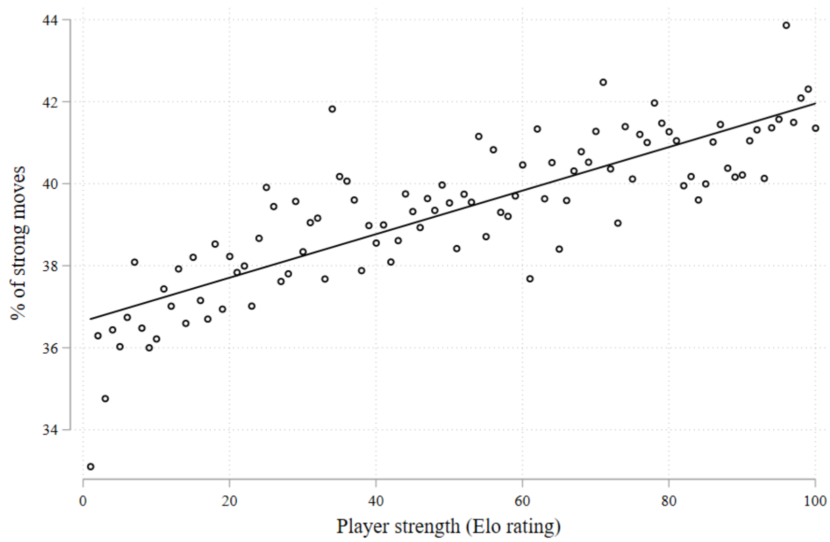



Figure 2.4: Mean percentage of strong moves per game by player strength

*Notes:* Elo rating in x-axis are percentiles of the Bayesian Elo measure of player strength at game day. Percentage of strong moves is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119 and is averaged at the Elo percentile level.

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

$$F\left[Pr(Black\ wins)\right] = \beta_0 + \beta_1 \mathbb{1}(\Delta strong > 0) + \beta_2 \mathbb{1}(\Delta Elo > 0) + \beta_3 \mathbb{1}(\Delta age > 0)\} \tag{2.2}$$

$$F\left[Pr(Black\ wins)\right] = \beta_0 + \beta_1 \mathbb{1}(\Delta inaccuracy > 0) + \beta_2 \mathbb{1}(\Delta Elo > 0) + \beta_3 \mathbb{1}(\Delta age > 0) \tag{2.3}$$

players rather than Elo ratings for measuring player strength. The weaker relationship between dan and the performance metrics among the strongest (high dan) players is likely attributed to dan representing a conferred degree and not the current strength of a player, meaning some players at the highest dan levels are potentially past their peak performance.

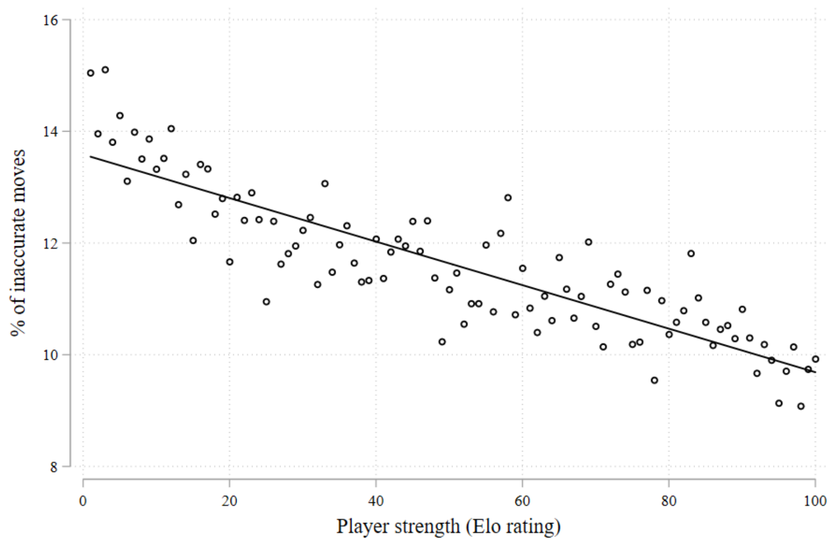



Figure 2.5: Mean percentage of inaccuracies per game across different ranks

*Notes:* Elo rating in x-axis are percentiles of the Bayesian Elo measure of player strength at game day. Percentage of inaccurate moves is calculated as the percentage of a player's move choices, in the mid-game range of moves 100-119, which are not in the set of candidate moves proposed by Leela Zero, averaged at the Elo percentile level.

where $F$ is the logit function linearizing the model.

$$F[x] = \ln\left[\frac{x}{1-x}\right] \tag{2.4}$$

In equation 2.2, $\mathbb{1}(\Delta strong > 0)$ equals one if the count of strong moves of Black player is higher than White's; similarly in equation 2.3 $\mathbb{1}(\Delta inaccurate > 0)$ equals one if Black makes more inaccurate moves than White. $\mathbb{1}(\Delta Elo > 0)$ takes value one if Black is higher rated than White; and, $\mathbb{1}(\Delta age > 0)$ equals one if Black is the older player.

Columns (1) to (3) of Table 2.3 present logistic regression results expressed in odds ratios for the model in equation 2.2, where the coefficients are the exponents of $\beta_1$, $\beta_2$ and $\beta_3$ 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 2.3. Both tables substantiate the claim that relative performance matters: the odds of winning are approximately 1.52 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

Table 2.3: Logistic Regression of Relative Performance on Game Outcome

| Dep. Var.: Pr(Black Wins) | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\Delta strong > 0$ | 1.499*** | 1.518*** | 1.519*** | | | |
| | (15.19) | (14.98) | (14.98) | | | |
| $\Delta inaccurate > 0$ | | | | 0.337*** | 0.350*** | 0.350*** |
| | | | | (-35.88) | (-33.37) | (-33.30) |
| $\Delta Elo > 0$ | | 3.347*** | 3.331*** | | 3.247*** | 3.231*** |
| | | (44.36) | (44.13) | | (42.41) | (42.18) |
| $\Delta age > 0$ | | | 0.801*** | | | 0.805*** |
| | | | (-8.11) | | | (-7.79) |
| Observations | 24066 | 24066 | 24066 | 24066 | 24066 | 24066 |

Exponentiated coefficients; $t$ statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* $\Delta strong$ and $\Delta inaccurate$ are the difference in count by the Black versus White player of strong moves and inaccurate moves, respectively. Similarly, $\Delta Elo$ is the Black versus White player difference in player strength as measured by Elo rating, and $\Delta age$ is the age difference of the Black versus White player. The coefficients shown are exponentiated and 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. $t$ statistics shown in parentheses.

0.35, and robust to the inclusion of covariates, for players making fewer inaccuracies. In plain English, these odds ratios translate to players performing relatively more strong moves winning 3 out of 5 games played, and players making more inaccuracies winning just 7 out of 27 games. Jointly, these table results and Figures 2.4 and 2.5 validate the move evaluations of Leela Zero as relevant measures of the quality of a player's decision-making in the game.

As previously mentioned, regressing the main specification in equation 2 only yields a causal effect under the assumption that the treatment variable of Asian dust events $Dust_{jt}$ is conditionally independent of the error term $\varepsilon_{pjt}$. One concern is that some Go players may engage in dust avoidance behaviour. In a first look into this possibility, Figure 2.6 plots histograms of games played each month during over a set of years where Asian dust received considerable media attention. 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

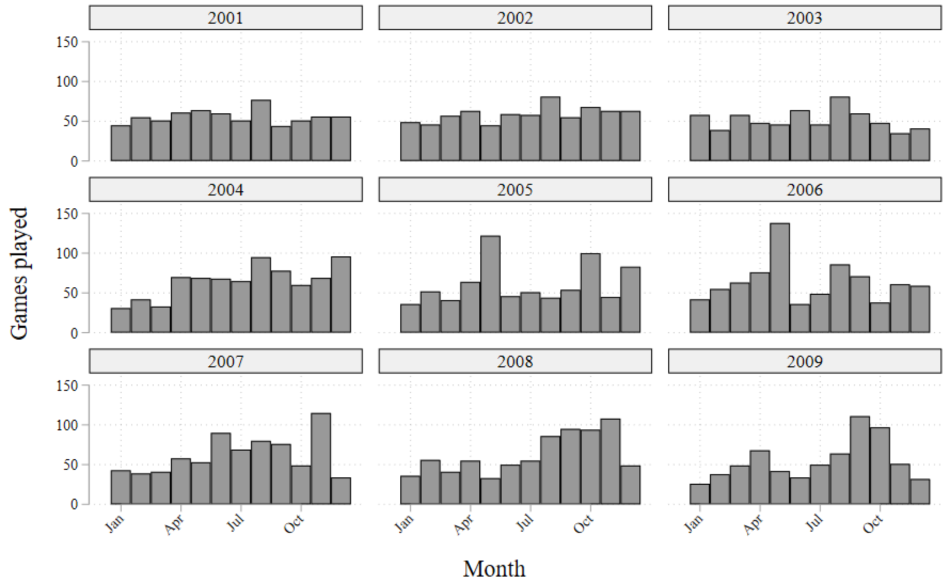



Figure 2.6: Histogram of games played each month over the years ranging 2001 and 2009

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 inflexible 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.

Stronger evidence against dust avoidance at the city-day level can be obtained with an event study variation of equation 2.1:

$$Y_{jt} = \alpha + \sum_{\tau} \delta_\tau Dust_{j,t+\tau} + \psi_j + \eta_{y(t)} + \eta_{m(t)} + \varepsilon_{jt} \tag{2.5}$$

where $Y_{jt}$ is the log number of games played in day $t$ and city $j$, and $\tau$ 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 days.[17] Figure 2.7 presents the estimated $\hat{\delta}_\tau$ coefficients and

[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.

standard errors for leads and lags $\tau$ around dust days. The observed pattern is suggestive that game scheduling is not sensitive to dust storm occurrences, strengthening the case against dust avoidance of players.[18]

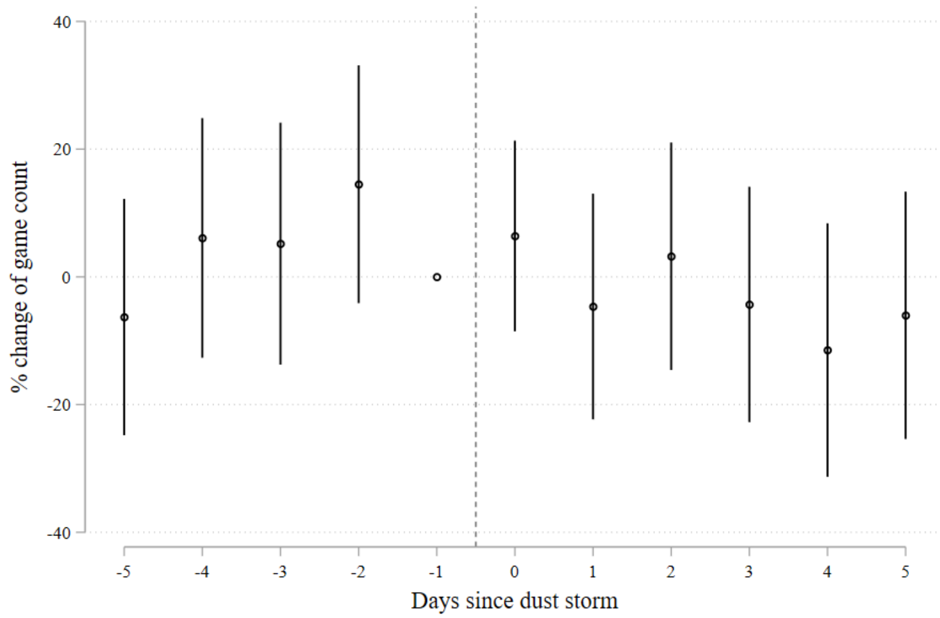



Figure 2.7: Change in number of games around dust event window

*Notes:* The y-axis % change is $100 * \hat{\delta}$, i.e., the percentage change in game counts $\tau$ days since a dust event. Sample includes only games within a 5-day window around Asian dust storms. Vertical lines correspond to 95% confidence intervals.

Next, I document the air pollution shock induced by Asian dust storms. Figure 2.8) presents event study estimates from the same model as in equation 5 with the log of an air pollutant concentration as outcome variable. The figure shows a sharp and short-lived increase in daily average concentration of PM pollutants (left panels), approximately 45% for $PM_{2.5}$, 75% for $PM_{10}$ and 35% for $SPM$, during the day of an Asian dust. The air pollution shock is qualitatively similar but smaller in magnitude for $O3$ and of even smaller magnitudes and in the opposite direction for $SO_2$ and $CO$. The puzzling results of smaller concentrations for certain pollutants following the Asian dust events could be explained by higher wind speeds during Asian dust events helping to eliminate local

[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.

build-up of man-made pollution that occurs when the air is stagnate (Yang et al., 2017).

The importance of these shocks also differs 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 $PM_{10}$ cleaner recent annual average of $22.8\mu g/m^3$ in 2016 to its dirtier annual average of $68.8\mu g/m^3$ in 2001. For $PM_{2.5}$, the 45% increase corresponds to an additional $15\mu g/m^3$ in pollutant concentration. The shocks diminish in importance for the other pollutants examined. For example, the 15% increase in $O_3$ during dust days in Figure 2.8 would imply an almost 3 parts per billion (ppb) concentration increase in major South Korean cities, much 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. The event study results overall point to PM pollution as the main driver of the equation 2 results presented below.[20]

## Overall impacts

The results up to this point support interpreting equation 2.1 as estimating a *causal* relationship of air pollution on the quality of decision-making of Go players. Tables 2.4 and 2.5 report the estimated impact of Asian dust storms on the constructed measures of Go players' cognitive performance. The tables are formatted to include additional controls and/or fixed effects at each subsequent column to the right. Column (1) includes time fixed effects and controls for player sex, column (2) then incorporates city fixed effects, column (3) controls for player strength measured by the Elo rating, and the richest and preferred specification in column (4) adds player fixed effects and drops the female dummy since it is a time-invariant player characteristic.

Table 2.4 reports the estimation results with the percentage of strong moves per game as dependent variable. I find small and statistically insignificant point estimates across all specifications, which would suggest that the quality of decision making is not significantly affected by Asian dust exposure.

[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.

[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.

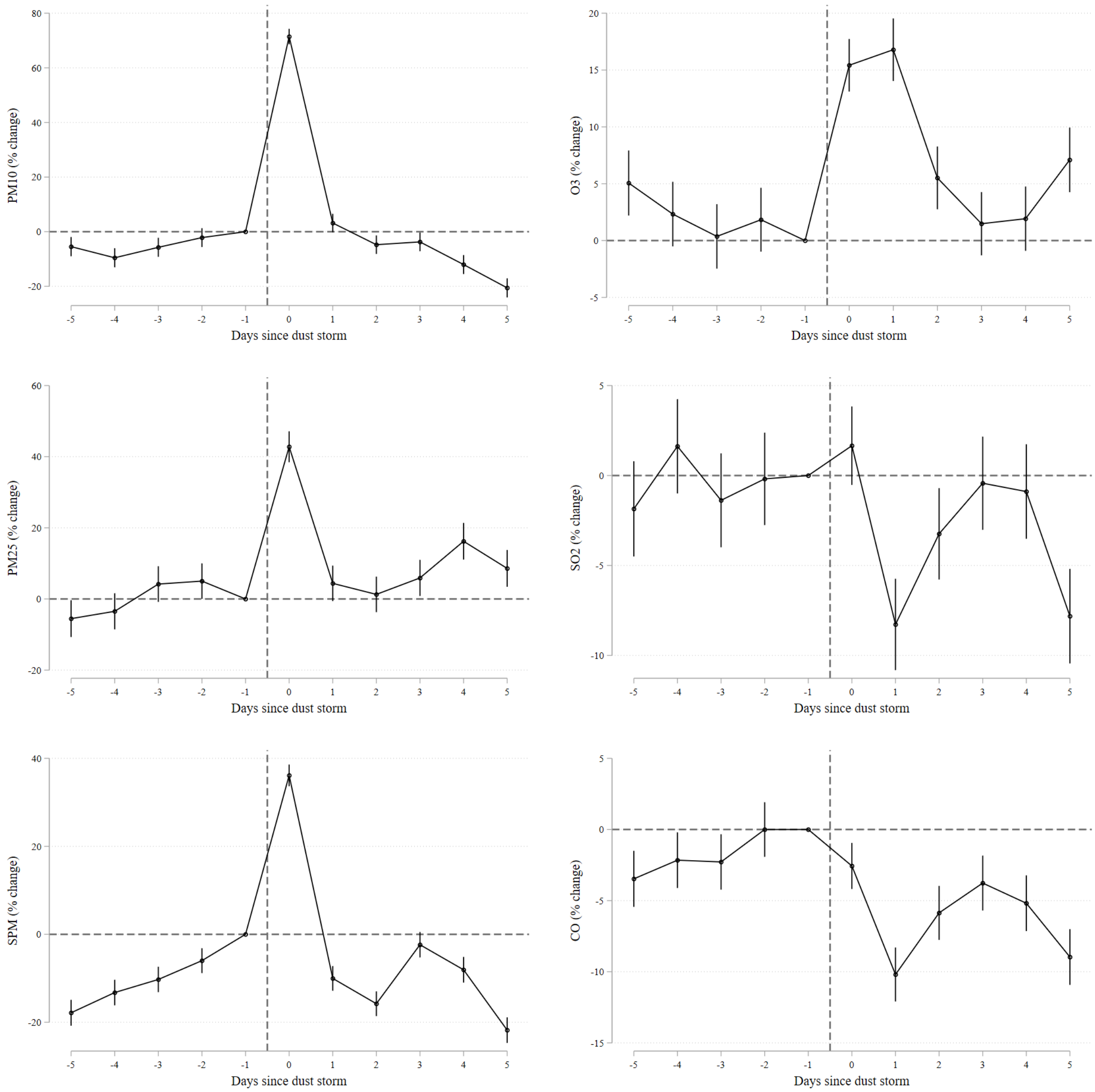



Figure 2.8: Changes in pollutant levels on five-day window before/after dust events

*Notes:* 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. Sample includes only air pollution observations within a 5-day window around Asian dust storms. Vertical lines correspond to 95% confidence interval.

Table 2.5 reports the results using the percentage of inaccuracies per game as dependent variable, in which case the estimates are suggestive that player moves become more inaccurate under exposure to Asian dust. The coefficients are reasonably stable across specification, and the preferred estimate

Table 2.4: Effect of Asian dust on percent of strong moves per game

| Dep. Var: %strong | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dust event | -0.052 | -0.024 | 0.012 | -0.065 |
| | (0.471) | (0.523) | (0.548) | (0.663) |
| Female=1 | -3.275*** | -3.227*** | -1.043** | |
| | (0.329) | (0.319) | (0.482) | |
| Player Elo rating | | | 0.008*** | 0.008*** |
| | | | (0.001) | (0.002) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| Observations | 44012 | 44005 | 44005 | 43880 |
| $R^2$ | 0.016 | 0.023 | 0.027 | 0.052 |

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %strong is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119.

implies that Go players make .967 percentage point more inaccuracies when exposed to Asian dust. This translates to an increase in inaccurate moves of 8.3% relative to the baseline average of 11.69% inaccuracies in days without dust events, or roughly one additional inaccuracy per one hundred moves. The findings from Tables 2.4 and 2.5 point towards an increase in the likelihood of making a human error due to the shock in air pollution exposure. The coefficients on gender in the regressions also reveal a bias from gender disparity in a game where there have been historically few amateur and even fewer professional female players. This selection bias reduces after controlling for player strength in column (3), where in these and subsequent regression tables the gender coefficients become attenuated or even switch sign.

Are all individuals in the sample equally susceptible to the effects of air pollution? I now turn to

Table 2.5: Effect of Asian dust on percent of inaccuracies per game

| Dep. Var: %inaccurate | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dust event | 1.170*** | 1.178*** | 1.145*** | 0.967** |
| | (0.341) | (0.338) | (0.367) | (0.442) |
| Female=1 | 1.589*** | 1.650*** | -0.352 | |
| | (0.160) | (0.169) | (0.319) | |
| Player Elo rating | | | -0.008*** | -0.005*** |
| | | | (0.000) | (0.001) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| Observations | 44012 | 44005 | 44005 | 43880 |
| $R^2$ | 0.013 | 0.017 | 0.027 | 0.060 |

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %inaccurate is calculated as the percentage of a player's move choices which are not in the set of candidate moves proposed by Leela Zero.

exploring treatment effect heterogeneity across demographic groups. The wide distribution of age and skill allow me to reproduce the main results over players differing in these two dimensions.[21]

## Impacts across age groups

Table 2.6 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 2.4 results. The coefficients are again statistically insignificant, with precisely estimated point estimates close to zero. The estimates become more negative in Panel B – which presents the estimation for players

[21] A stratification by gender is not feasible in my analysis as games with female players are rare occurrences in the sample.

above median age – however remain insignificant even at the 10% significance level.

Table 2.7 likewise reports the results by age group with the percentage of inaccurate moves per game as dependent variable. 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 1.78pp more inaccurate moves per game on average using the preferred estimate from column (4). This estimate implies an increase in human error of 14.7% for the older players or close to two additional inaccuracies per one hundred moves played. Compared to Table 2.5, these results are strongly indicative of heterogeneity, with older players being more susceptible to deterioration of cognitive performance on dust days. This result confirms evidence from the literature of age heterogeneous effects of pollution exposure on cognitive functioning such as Lai et al. (2022) and Krebs and Luechinger (2021).

### Impacts across player ranks

Figures 2.4 and 2.5 have shown a strong relationship between the cognitive performance outcomes and player strength measured by Elo ratings. In light of this, I stratify the players in the sample by their dan degree levels to investigate whether the cognitive performance of players in different ranks responds differently to the pollution shocks induced by Asian dust storm. I separate the sample into the stronger high dan, comprising players who were awarded a dan degree of 5d to 9d, and low/amateur dan comprising players with dan degrees below 5d. Table 2.8 reports the output from regression equation 2.1 using the percentage of strong moves outcome. Panel A shows coefficient estimates for the weaker players become generally positive (although far from significant on any specification), which would suggest air pollution improves the quality of decisions for these individuals. The output of panel B is very similar to the output for the overall sample which is skewed towards high dan individuals. The coefficients are negative, statistically insignificant, and fairly stable across specifications, suggesting as before little relationship between air quality and this margin of decision-making quality.

Table 2.9 reports the estimates using the percentage of inaccuracies outcome. The effects of air pollution on inaccuracies are visible and significant for both stronger and weaker players. Low and amateur dan players suffer the largest loss in cognitive performance, with inaccuracies moves increasing 1.88 percentage points, or nearly two additional inaccuracies per one hundred moves). Inaccuracies by high-dan players increase 1 percentage point, implying as in the overall sample on average one additional inaccuracy per one hundred moves. Although the low/amateur dan players have a higher baseline likelihood of playing inaccuracies, the estimated magnitudes in this table suggest these weaker players are more adversely impacted by Asian dust. A possible explanation for these differing results is that individuals at the top of Go rankings are highly selected and may exhibit certain traits – such as constancy and higher concentration – making them less susceptible to cognitive losses induced by air pollution. Future work could identify how elite Go players differ from the general population.

## 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 2.10, shows coefficient estimates as well as patterns of statistical significance which are fairly consistent with the results from previous subsections.

Table 2.6: Effect of Asian dust on percent of strong moves per game by age group

| Dep. Var: %strong | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Panel A** | | | | |
| **Below median age (30 yrs)** | | | | |
| Dust event | -0.161 | 0.048 | -0.023 | -0.213 |
| | (0.593) | (0.680) | (0.698) | (0.920) |
| Female=1 | -3.335*** | -3.244*** | -1.618*** | |
| | (0.248) | (0.248) | (0.298) | |
| Player Elo rating | | | 0.006*** | 0.007 |
| | | | (0.001) | (0.005) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.031 | 0.039 | 0.040 | 0.075 |
| Observations | 21344 | 21308 | 21308 | 21217 |
| | (1) | (2) | (3) | (4) |
| **Panel B** | | | | |
| **Above median age (30 yrs)** | | | | |
| Dust event | -0.968 | -1.198 | -0.954 | -1.151 |
| | (0.808) | (0.802) | (0.815) | (0.785) |
| Female=1 | -3.663*** | -3.614*** | -1.043 | |
| | (0.765) | (0.751) | (0.975) | |
| Player Elo rating | | | 0.010*** | 0.001 |
| | | | (0.001) | (0.002) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.026 | 0.035 | 0.040 | 0.069 |
| Observations | 21807 | 21775 | 21775 | 21649 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %strong is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119.

Table 2.7: Effect of Asian dust on percent of inaccuracies per game by age group

| Dep. Var: %inaccurate | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Panel A** | | | | |
| **Below median age (30 yrs)** | | | | |
| Dust event | 0.751 | 0.804 | 0.878 | 0.717 |
| | (0.787) | (0.756) | (0.784) | (0.983) |
| Female=1 | 1.581*** | 1.608*** | -0.063 | |
| | (0.250) | (0.247) | (0.301) | |
| Player Elo rating | | | -0.006*** | -0.005*** |
| | | | (0.001) | (0.002) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.025 | 0.031 | 0.036 | 0.078 |
| Observations | 21344 | 21308 | 21308 | 21217 |
| | (1) | (2) | (3) | (4) |
| **Panel B** | | | | |
| **Above median age (30 yrs)** | | | | |
| Dust event | 2.129*** | 2.120*** | 1.887*** | 1.783*** |
| | (0.536) | (0.538) | (0.475) | (0.602) |
| Female=1 | 1.523*** | 1.644*** | -0.815** | |
| | (0.476) | (0.472) | (0.402) | |
| Player Elo rating | | | -0.009*** | -0.003*** |
| | | | (0.000) | (0.000) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.024 | 0.030 | 0.044 | 0.080 |
| Observations | 21807 | 21775 | 21775 | 21649 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %inaccurate is calculated as the percentage of a player's move choices which are not in the set of candidate moves proposed by Leela Zero.

Table 2.8: Effect of Asian dust on percent of strong moves per game by player strength

| Dep. Var: %strong | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Panel A** | | | | |
| **Low/Amateur Dan** | | | | |
| Dust event | 0.997 | 0.766 | 0.755 | 0.108 |
| | (0.923) | (0.735) | (0.755) | (0.756) |
| Female=1 | -2.944*** | -2.894*** | -1.264* | |
| | (0.625) | (0.614) | (0.745) | |
| Player Elo rating | | | 0.008*** | 0.011** |
| | | | (0.001) | (0.005) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.060 | 0.069 | 0.071 | 0.165 |
| Observations | 8442 | 8420 | 8420 | 8268 |
| | (1) | (2) | (3) | (4) |
| **Panel B** | | | | |
| **High Dan** | | | | |
| Dust event | -0.287 | -0.287 | -0.229 | -0.318 |
| | (0.647) | (0.687) | (0.741) | (0.784) |
| Female=1 | -2.967*** | -2.908*** | -0.947 | |
| | (0.487) | (0.484) | (0.612) | |
| Player Elo rating | | | 0.008*** | 0.005** |
| | | | (0.000) | (0.002) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.018 | 0.026 | 0.029 | 0.050 |
| Observations | 35284 | 35277 | 35277 | 35192 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %strong is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119.

Table 2.9: Effect of Asian dust on percent of inaccuracies per game by player strength

| Dep. Var: %inaccurate | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Panel A** | | | | |
| **Low/Amateur Dan** | | | | |
| Dust event | 1.163 | 1.505** | 1.517** | 1.882*** |
| | (0.932) | (0.737) | (0.669) | (0.698) |
| Female=1 | 1.179*** | 1.204** | -0.483 | |
| | (0.441) | (0.515) | (0.497) | |
| Player Elo rating | | | -0.008*** | -0.009*** |
| | | | (0.001) | (0.002) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.056 | 0.061 | 0.069 | 0.166 |
| Observations | 8442 | 8420 | 8420 | 8268 |
| | (1) | (2) | (3) | (4) |
| **Panel B** | | | | |
| **High Dan** | | | | |
| Dust event | 1.259*** | 1.289*** | 1.236*** | 0.998*** |
| | (0.258) | (0.269) | (0.270) | (0.345) |
| Female=1 | 1.421*** | 1.564*** | -0.230 | |
| | (0.340) | (0.346) | (0.355) | |
| Player Elo rating | | | -0.008*** | -0.004*** |
| | | | (0.000) | (0.001) |
| Year & month FE | ✓ | ✓ | ✓ | ✓ |
| City FE | | ✓ | ✓ | ✓ |
| Player FE | | | | ✓ |
| $R^2$ | 0.015 | 0.020 | 0.029 | 0.057 |
| Observations | 35284 | 35277 | 35277 | 35192 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %inaccurate is calculated as the percentage of a player's move choices which are not in the set of candidate moves proposed by Leela Zero.

Table 2.10: 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 | -0.240 | -0.878 | -1.112 | -0.834 | -0.470 |
| | (0.699) | (0.819) | (0.836) | (0.530) | (0.833) |
| %inaccurate | $0.871^{**}$ | 0.842 | $1.435^{**}$ | $2.660^{***}$ | $0.849^{***}$ |
| | (0.385) | (0.830) | (0.610) | (0.624) | (0.297) |
| Observations | 28878 | 13027 | 15493 | 3268 | 25479 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes:* Years covered in the regression are 1980 to 2017. Regressions in these table reproduce the preferred specification for Tables 2.4 to 2.9 after excluding from the sample the players who did not play any game in an Asian dust day. A player's move is defined as strong if it coincides with the move achieving the highest VN score among the candidate moves evaluated by the AI. %strong is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119. %inaccurate is calculated as the percentage of a player's move choices which are not in the set of candidate moves proposed by Leela Zero.

## 2.7 Conclusion

I have examined 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 pollutants.

I also constructed 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 evaluated moves played by expert Go players and classified them as strong, acceptable, or inaccurate. 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 make inaccurate moves 8.3% more often. In absolute terms, this amounts to roughly one additional inaccuracy per one hundred moves, amounting to two additional inaccuracies over an average length game of 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 make nearly two additional inaccuracies per one hundred moves.

For a comparison: Archsmith et al. (2018) estimate that Major League of Baseball umpires make one additional ball/strike incorrect call for every 250 decisions when $PM_{2.5}$ levels increase $10\mu g/m^3$. A back-of-the-envelope calculation, using the average $15\mu g/m^3$ increase in $PM_{2.5}$ during Asian dust days and the baseline effect on inaccuracies of $\delta = 0.967$, implies my estimated effect is 61% larger in magnitude than the effect on incorrect calls of baseball umpires. Similarly, the findings from Ebenstein et al. (2016) suggest a $10\mu g/m^3$ increase in $PM_{2.5}$ raises the probability of a student failing the high-stakes Bagrut test by 2 percentage points. [22] This magnitude is slightly higher than the strongest adverse effect in my analysis, a 1.88 percentage point increase in inaccuracies by the low/amateur dan players.

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. Prior studies 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.

[22] The authors report air pollution concentration using an air quality index, but note that this index has a correlation of .9855 with cubic micrograms ($\mu g/m^3$).

# Chapter 3

# Policy-Induced Innovation in Energy Storage

## 3.1 Introduction

Renewable resources play an important role in powering cities around the world. And while renewable sources such as wind and solar support policy goals of lowering greenhouse gas (GHG) emissions, their intermittent nature poses new challenges to the power grid. (Gowrisankaran et al., 2016). Today, energy storage technologies represent an increasingly viable option for improving the reliability of power grids relying on solar and other renewable sources. In the United States, the importance of energy storage can be recognized by policy adoption, regulations, and market design (Sakti et al., 2018).

Complementing the policy side of energy storage, innovation research is key to unlocking the transformational potential of storage technologies. Policymakers form expectations about future innovation when making decisions on subsidies for technology push and demand pull. Similarly, firms and consumers draw on such information for generating and adopting new technologies respectively. Because the economic potential for new technologies is large, novel ideas are likely first patented before scientifically published, if at all (Mueller et al., 2015).

I connect the literature on energy storage policy and innovation by considering the role between governmental policies affecting investments in energy storage and future patenting activity in these energy storage technologies. My goal is to shed light on how patenting activity in energy storage is related to changes in the market size for these technologies, and subsequently to incorporate government policies as a determinant of market size. I observe changes in the energy sector induced by a new regulation in the market for frequency regulation, an ancillary service essential for grid reliability. Because various energy storage technologies (e.g. batteries and flywheels) provide accurate and fast responses to changes in electricity demand, new market rules for frequency regulation led to greater investment incentives which in turn caused an increase in the number of energy storage projects capable to provide these ancillary services. The regulatory change in question was Order 755, passed in 2011 by the Federal Energy Regulatory Commission (FERC), a federal agency regulating interstate transmission and wholesale electricity markets in the US.

Order 755 directed independent system operators and regional transmission operators[1] to compensate frequency regulator providers under the FERC jurisdiction for improving the service's speed and accuracy. Previous research shows that US regions under FERC jurisdiction saw increases in the likelihood of projects targeting frequency regulation, and more importantly in the yearly number and installed capacity of energy storage projects (Tabari and Shaffer, 2020).

To measure innovation in energy storage, I build a novel data set that classifies granted US patent data between 2000 to 2018 from PatentsView based on the presence of energy storage technology concepts. To do so, I construct queries using Dimensions Search Language (DSL), a language used to retrieve information from the Dimensions Patents dataset. These queries allow me to map energy storage concepts to Cooperative Patent Classification (CPC) codes available in PatentsView, aggregated at the main group level, and identify energy storage-related patents among all US patent grants. The queried energy storage concepts are sourced from technology classifications of energy storage projects used by the Global Energy Storage Database (GESDB), an archive of energy storage projects maintained by the US Department of Energy (DOE) with information on announcement dates, location, project capacity, technology type, and more. The technology classification of projects in the GESDB has been carefully curated based on a taxonomy published in the DOE's Energy Storage Handbook (Nguyen et al., 2021). By using technology types from the GESDB to construct this new classification of energy storage technologies, I ensure the definition of energy

[1]Loosely defined, these are organizations ensuring the reliable and efficient operation of the electricity grid.

storage technologies in the analysis is closely aligned with the energy storage market size, which is proxied with the deployment of storage projects observed in the GESDB.

I use a two-way fixed effects model to explore the relationship between the deployment of energy storage projects in a county and local patenting activity. The arrival of a new energy storage project is associated with an increase of 2.76% in the county's number of patents in energy storage technologies, but no effect on other (not storage-related) patents. Using an event study, I find this increase in patenting is short-lived and contemporaneous to the announcement of the storage project, suggesting a limited role of these storage projects on local innovation.

I then connect these findings with government policies affecting market conditions for energy storage. To do so, I employ a difference-in-differences approach to identify the effect of the FERC Order 755 in 2011 on energy storage patenting activity relative to non-storage closely related technologies. This policy change increases the compensation received by fast-acting frequency regulation resources, a market in which energy storage technologies can contribute greatly. I find that the introduction of Order 755 led to more patents in energy storage relative to the closely related technologies, suggesting a role of policy on guiding the direction of innovation.

This paper contributes to our understanding about the role of energy policy on innovation, where prior literature has focused on energy prices (Newell et al., 1999; Popp, 2002; Acemoglu et al., 2012). Other strand of the literature emphasizes the importance of environmental policy more broadly, such as the work of Aghion et al. (2016) who provide evidence that carbon taxes promote innovation in clean technologies at the expense of less innovation in dirty technologies in the automobile industry.

Since my emphasis is on policies affecting market size, Acemoglu et al. (2012) is particularly relevant for my context, as they incorporate market size along with energy prices in their framework to study induced innovation in the energy sector. In recent work, Feng and Lazkano (2022) highlight both the importance of energy prices and past innovation in shaping the direction of current innovation in energy storage. Ultimately, the goal of my work, along with this related literature, is to contribute to the toolkit of policymakers on promoting innovation. [2]

[2] Bloom et al. (2019) offers an excellent review of the available policy levers "aimed at increasing the supply of human capital focused on innovation, intellectual property policies, and pro-competitive policies."

The paper is organized as follows. In the next section I describe the data on patents and on energy storage projects. Section 3 presents the empirical strategy and reports findings on the relationship between energy storage market size and innovation. Section 4 presents a different empirical strategy and reports findings on the subsequent analysis on the role of energy storage policy in innovation. Section 5 concludes the paper.

## 3.2 Data

Data on granted US patents between 2000 and 2020 was sourced from PatentsView[3]. The dataset includes, for each patent, the filing date, county of each listed inventor, a comprehensive list of current CPC codes (aggregated at the main group level) assigned to the patent and current citations from granted US patents[4]. The main patent sample comprises 2,492,489 granted patents, with filing dates between Jan. 1, 2000 and Dec. 31, 2020, and for which at least one inventor resides in the US.

Additionally, data on energy storage projects come from the GESDB and includes project announcement date, location, technology type, and more. The sample consists of 471 energy storage projects for which the announcement date is available. In the analysis, the project year refers to the announcement date[5]. These projects span the years 2000 to 2018. Nine pumped hydro storage projects are dropped from the sample[6]. Among the 471 energy storage projects in the sample, 116 do not list the firm contracted for providing the storage technology. From the remaining 355, 309 (87%) report a firm containing US patent records. These 309 firms include well-established companies such as Samsung SDI, LG Chem, and Tesla Energy, as well as energy storage solutions startups. The geographical distribution of projects is shown in the map in Figure 3.1.

As the primary interest is on patents related to energy storage, I treat the technology types reported in the GESDB as a collection of energy storage concepts. I query these concepts on the title

[3]Specifically, I download the February 21, 2023 snapshot of PatentsView.

[4]CPC codes have changed over time, and the codes reported in the data correspond to those assigned at the time the data was extracted. Similarly, patent citations are only available from US granted patents prior to data extraction.

[5]Construction/commissioning dates are known for only 157 of the 471 projects in the sample. In addition, more than half of the projects where only announcement date is known are listed as operational.

[6]Pumped hydro storage projects are disproportionately large (in size and capacity) and are all in preliminary development stages.

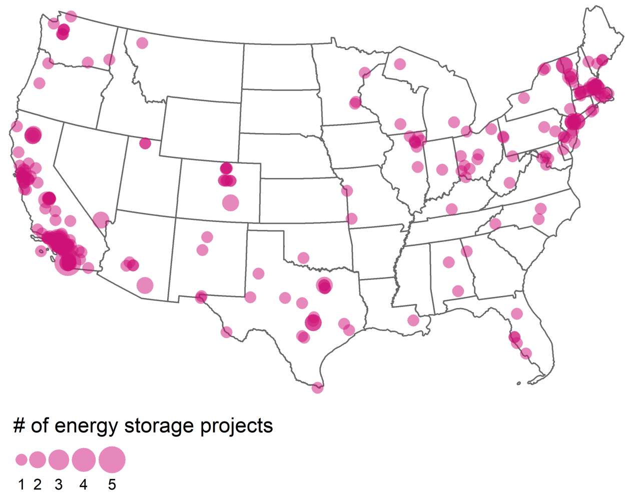



Figure 3.1: Geographical distribution of energy storage projects in GESDB

and description/summary text of patents using the DSL API of the Dimensions dataset and collect the CPC codes at the main group level reported on the patents matching the energy storage concepts. I select the ten most frequent CPC codes for each energy storage concept and compile a list of CPC codes related to energy storage technologies. Patents with a primary CPC code belonging to this set are classified as energy storage-related, with the additional requirement that the patents also list a secondary CPC code from the Y02E60 main group, which identifies enabling technologies contributing to GHG emissions mitigation. Panel (a) of Figure 3.2 shows the distribution of types of technology in the energy storage projects found in the main sample from the GESDB. Similarly, panel (b) shows the number of patents matched when searching each energy storage concept against the patent title and text using the DSL API. The distribution across the two panels is similar for some technologies but not for others, possibly highlighting the gap in technology maturity between innovation and project deployment.

Table 3.1 shows the ten most cited patents classified in three broad technology types, where the electro-chemical category includes all battery storage (e.g. lithium-ion, flow batteries) and the electro-mechanical category includes flywheels and compressed air energy storage (CAES). In most applications in the GESDB and patent data, thermal storage refers to the process of chilling

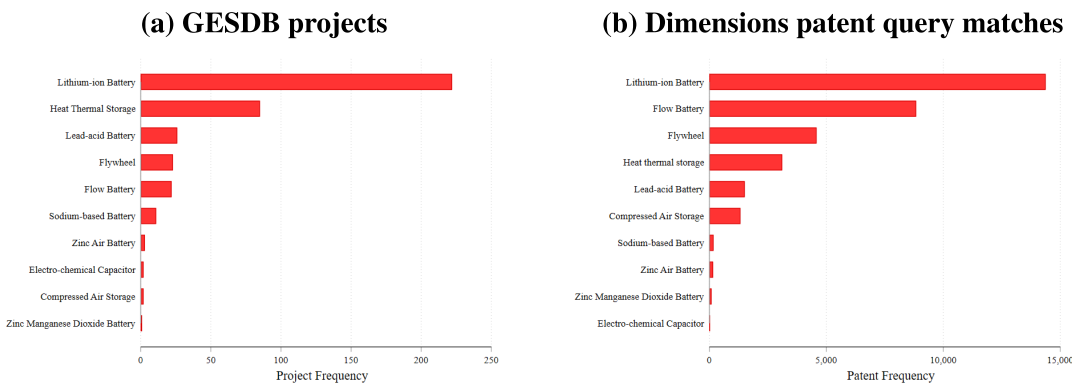



Figure 3.2: Technology concepts of energy storage

*Notes:* 74 energy storage projects in the sample report a broader electro-chemical technology, leaving the technology type unspecified. These projects are included in the analysis, but not shown in panel (a).

or freezing and using that thermal load to cool buildings and offices. The table illustrates that the innovations captured by this classification procedure are indeed related to energy storage, although not necessarily related to grid-scale applications.

Table 3.1: Most frequently cited patents, by broad technology types

| Patent Number | Publication Date | Forward Citations | Title of Patent |
|---|---|---|---|
| *Electro-chemical* | | | |
| US9153994 | 12oct2012 | 417 | Motion sensitive and capacitor powered handheld device |
| US9030169 | 03mar2009 | 369 | Battery system and method for system state of charge determination |
| US9722236 | 15mar2013 | 241 | Apparatus and method for use in storing energy |
| US8575895 | 29mar2011 | 209 | Method and device for voltage detection and charging of electric battery |
| US9337668 | 27apr2012 | 209 | Viral distribution of battery management parameters |
| US8487480 | 16dec2009 | 203 | Wireless energy transfer resonator kit |
| US8367235 | 21jan2009 | 166 | Battery pack, holster, and extendible processing and interface platform for mobile devices |
| US8350526 | 25jul2011 | 115 | Station for rapidly charging an electric vehicle battery |
| US8508188 | 15apr2010 | 100 | Universal charge module |
| US8482255 | 26aug2010 | 99 | Method and system for charging an auxilary battery in a plug-in electric vehicle |
| *Electro-mechanical* | | | |
| US7802426 | 09jun2009 | 91 | System and method for rapid isothermal gas expansion and compression for energy storage |
| US8046990 | 14feb2011 | 56 | Systems and methods for improving drivetrain efficiency for compressed gas energy storage |
| US8104274 | 18may2011 | 31 | Increased power in compressed-gas energy storage and recovery |
| US8250863 | 27apr2011 | 24 | Heat exchange with compressed gas in energy-storage systems |
| US8011189 | 08dec2009 | 22 | Retrofit of simple cycle gas turbine for compressed air energy storage application having... |
| US7669423 | 26jan2009 | 22 | Operating method for CAES plant using humidified air in a bottoming cycle expander |
| US8171728 | 08apr2011 | 22 | High-efficiency liquid heat exchange in compressed-gas energy storage systems |
| US8245508 | 15apr2011 | 18 | Improving efficiency of liquid heat exchange in compressed-gas energy storage systems |
| US8272212 | 11nov2011 | 18 | Systems and methods for optimizing thermal efficiencey of a compressed air energy storage |
| US8341964 | 27oct2009 | 16 | System and method of using a compressed air storage system with a gas turbine |
| *Heat thermal* | | | |
| US9151545 | 20dec2010 | 27 | Thermal management of an electrochemical cell by a combination of heat transfer fluid... |
| US10197338 | 22aug2013 | 24 | Building system for cascading flows of matter and energy |
| US8722222 | 10jul2012 | 23 | Thermoelectric-based thermal management of electrical devices |
| US7938989 | 20oct2009 | 18 | Composite structures for storing thermal energy |
| US8769972 | 25nov2009 | 11 | Electrochemical compressor and refrigeration system |
| US8187731 | 05may2010 | 11 | Metal ferrite spinel energy storage devices and methods for making and using same |
| US8851066 | 01apr2010 | 9 | Thermal energy storage system |
| US8116913 | 20apr2009 | 8 | Heating and cooling system using compressed fluid |
| US7905110 | 02apr2009 | 8 | Thermal energy module |
| US7827807 | 06feb2009 | 8 | Refrigerant-based thermal energy storage and cooling system with enhanced heat exchange... |

*Notes:* The electro-chemical category combines patents with a primary CPC code related to lithium-ion batteries, lead-acid batteries, flow batteries, sodium-based batteries, zinc air batteries, electro-chemical capacitors, and zinc manganese dioxide batteries. Similarly the electro-mechanical category refers to flywheel and compressed air storage.

## 3.3 Energy Storage Market Size and Innovation

I first examine the association between new energy storage projects – a quantity related to the size of energy storage markets – and local innovation at the county level. From the PatentsView data, I construct a county-by-year panel of patent counts. Because patents potentially have multiple inventors residing in different places, I divide each patent by the number of inventors, assign the corresponding patent fraction to the inventor county, and aggregate these patent fractions for each county. I use patent filing date – as opposed to grant date – since this better captures the timing of innovation. To understand the relationship between market size and innovation in energy storage, I use the following two-way fixed effects estimation equation:

$$Y_{jt} = \beta_1 NewProjects_{jt} + \gamma_j + \theta_{s(j),t} + \varepsilon_{jt}, \tag{3.1}$$

where $Y_{jt}$ is the log of the number of patents in county $j$ and year $t$. $NewProjects_{jt}$ is a count of the number of new energy storage projects in a particular county and year This main specification also includes county fixed effects $\gamma_j$ and state-by-year fixed effects $\theta_{s(j),t}$ (where $s(j)$ maps counties into states) to control for time-invariant determinants of innovation that are specific to counties and differential evolution in patenting trends in different US states. Errors are clustered at the county level. Equation 3.1 is estimated separately using the count of patents in energy storage technologies and the count of all patents, of which energy storage-related patents make less than 3%.

The parameter $\beta_1$ represents the percentage change in the number of patents as an additional energy storage project is announced on the county. I assess how robustness of this estimated effect to changes in the main specification, and examine the pre-trends graphically with an event study to understand whether there is any anticipatory effect.

Table 3.2 shows the results from estimating equation 3.1. We go from a more parsimonious model on the leftmost column, with county and year fixed effects in column (1) to the preferred specification allowing different trends in patenting outcome at the state level in column (3). The preferred estimate suggests new energy storage projects in a county are associated with an increase of 2.76% the number of energy storage patents. The effect is contained within energy storage innovation, as columns (4) to (6) show a tightly near zero coefficient from estimating the model using the count of all other (non-storage related) patents as dependent variable. As patent outcomes are zeros for several counties, I check in Appendix Table C1 the robustness of these results to using the inverse hyperbolic sine (IHS) transformation to the patent counts, which has of being well-defined for zero values.[7] For larger values, the IHS transformation yields an output similar to the log transformation.

In a variant of equation 3.1, I introduce leads and lags to consider how patenting outcomes relate over time to new energy storage projects. This event study specification follows the estimation

[7]For an outcome x, the IHS is defined as $IHS(x) = ln(x + \sqrt{x^2 + 1})$.

Table 3.2: Effect of new energy storage projects on patenting

| Dependent Variable = | $Log$(Energy storage patents) | | | $Log$(All other patents) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $NewProjects$ | 0.03337*** | 0.03163*** | 0.02763** | 0.00299 | -0.00386 | -0.00346 |
| | (0.00952) | (0.01214) | (0.01196) | (0.00429) | (0.00484) | (0.00528) |
| County fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | | | ✓ | | |
| Year × ISO/RTO fixed effects | | ✓ | | | ✓ | |
| Year × State fixed effects | | | ✓ | | | ✓ |
| Observations | 6,477 | 6,477 | 6,414 | 24,080 | 24,080 | 24,069 |
| Number of Clusters | 1,007 | 1,007 | 998 | 2,728 | 2,728 | 2,727 |
| R-Square | 0.813 | 0.817 | 0.830 | 0.921 | 0.921 | 0.923 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Estimates shown are based on equation 3.1. Errors are corrected for clustering at the county level.

equation:

$$Y_{jt} = \sum_{t+\tau} \beta_\tau NewProjects_{j,t+\tau} + \gamma_j + \theta_{s(j),t} + \varepsilon_{jt}, \tag{3.2}$$

where $NewProjects_{j,t+\tau}$ represents the count of new energy storage projects in the county in year $\tau$ relative to $t$. In the analysis, $\tau$ runs from $-4$ to 4 with $\tau = -1$ being omitted, and thus the set of coefficients $\beta_\tau$ represents the dynamic effect of new energy storage projects relative to $t - 1$, which is the year prior to new storage projects in the county.

The event study results from estimating equation 3.3 separately on the counts of energy storage-related and on the count of all other patents are shown graphically in Figure 3.3 with a 95% confidence interval around the point estimate. The figure shows no evidence of an anticipatory effect on the years preceding the announcement of the storage project. The contemporaneous effect at the time of announcement ($t = 0$) is positive and significant, around 3%, for energy storage-related patents, similar in magnitude to the estimated effect in Table 3.2. The effect is short-lived, being close to zero and insignificant in subsequent years, contrary to what would be expected if the deployment of energy storage projects generated a positive knowledge spillover effect that increased local innovation. I also do not find an effect on patents in other technologies, neither before nor after a new energy storage project, lending further support to the claim that counties receiving new storage projects are not experiencing any particular growth in innovation.

In technology-driven industries, such as energy storage, invention in its early stages are characterized by uncertainty about the technological effectiveness and manufacturing feasibility (Kim et al., 2016). The contemporaneous and short-lived effect from this estimation is consistent with a strategic decision of firms of early patenting as a commitment to an energy storage technology. In addition, patents represent an imperfect measure of innovative outputs, as has been highlighted by many authors in this literature (Griliches, 1998; Lanjouw et al., 1998). In particular, granted patents represent the outcome of many years of research and development. A limitation of the event study analysis is the possible truncation that occurs as knowledge spillovers may not yet have materialized for projects announced near the end of the sample period.

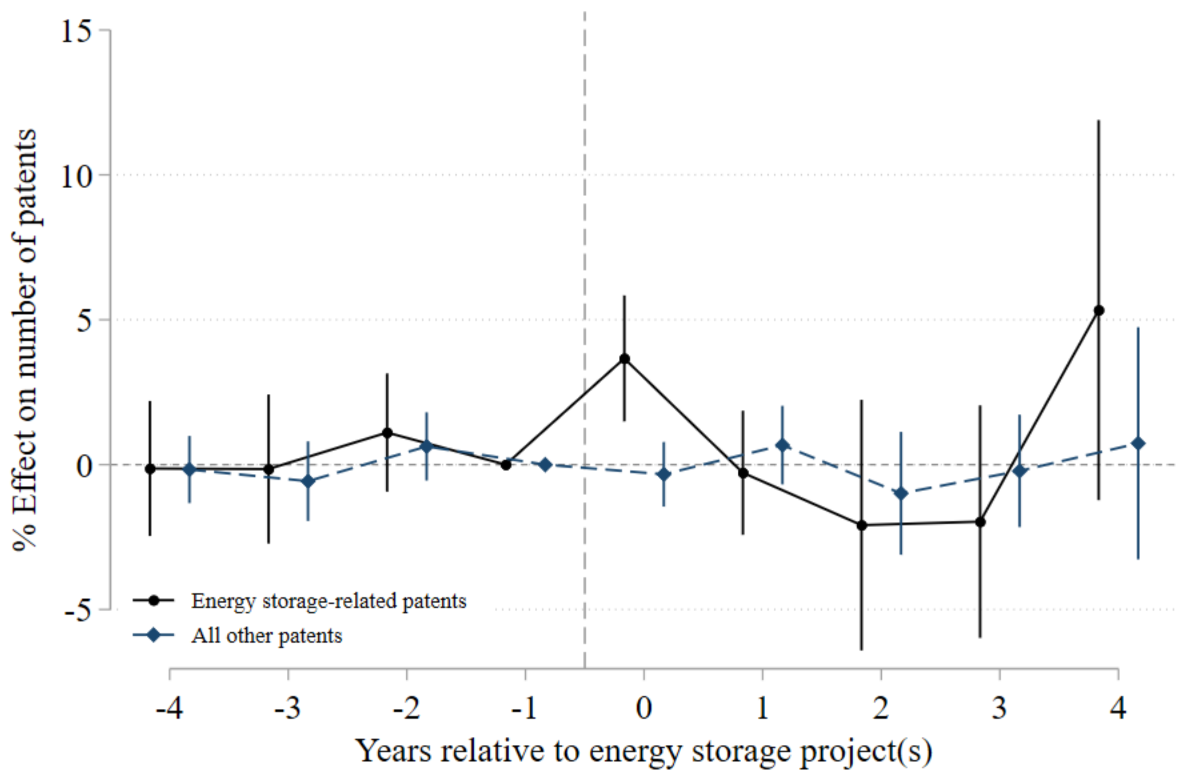



Figure 3.3: Event study: effect of new energy storage projects on patenting

*Notes:* Estimates shown are based on equation 3.2. Errors are corrected for clustering at the county level.

## 3.4 Energy Storage Policy and Innovation

The evidence presented in the previous section suggests that local innovation responds to changes in local market size induced by new energy storage projects. The timing of the estimated effects allows for an interpretation of firms committing to an energy storage technology prior to a project deployment. In contrast to my county-level evidence, prior work such as Gissey et al. (2018) suggests that energy storage innovation is driven by learning through deployment, and regulatory barriers often slow down innovation by hindering the market entry of storage technologies. My analysis so far exploits local variation in market size, where knowledge spillovers would be driven by geographical proximity between corporate R&D labs. However, these spillovers may be mediated over another dimension, such as technological distance. I now turn to exploring this possibility by conducting analysis at a national level.

In the context of the US energy storage markets, Tabari and Shaffer (2020) show that energy storage projects were responsive to FERC Order 755 – which incentivized energy storage systems to participate in the frequency regulation market. The targeted nature of this order offers a potential channel through which this particular policy could generate spillover effects in innovation complementary to the technologies deployed in these energy storage projects.

To investigate this possibility, I use an empirical approach for comparing patenting outcomes between a "treated" group of patents in energy storage – defined as in the previous section – versus a "control" group of non-storage patents from closely related technology classes. The control group of patents comprises non-storage patents for which the primary[8] CPC subclass (i.e., the first four digits of the CPC code) contains a (more granular) CPC main group tagged as energy storage-related in my classification procedure. For example, the CPC subclass H02J (circuit arrangements or systems for supplying or distributing electric power; systems for storing electric energy) which contains for instance the energy storage-related CPC main group H02J7 (circuit arrangements for charging or depolarising batteries or for supplying loads from batteries) also contains the non-storage main group H02J1 (circuit arrangements for dc mains or dc distribution networks). Patents with a primary

[8]The "primary" CPC is the classification that best describes the claimed invention. It is also the one appearing first on the published patent document.

CPC main group H02J1 integrate this control group of closely related technologies.

I aggregate the patent counts over CPC main groups of energy storage and closely related technology classes and over filing year. With this data, I employ a difference-in-differences (DiD) design comparing patenting outcomes over energy storage versus similar non-storage technologies, before versus after the introduction of Order 755 in October 2011. Patents from CPC main groups that are neither tagged as energy storage or closely related technologies are excluded from the analysis. The sample includes patents from 2006 to 2018, providing a six-year window around the event. I estimate the model with the regression specification:

$$PC_{ct} = \beta_1(StorageCPC_c \times Post_t) + \nu_c + \theta_t + \varepsilon_{ct}, \tag{3.3}$$

where $PC_{ct}$ is the count of all patents in the CPC main group $c$ in year $t$. $StorageCPC_c$ is an indicator equal to 1 if the CPC main group has been classified as energy storage-related and equal to 0 for the closely related technologies, and $Post_t$ is an indicator equal to 1 at or after 2012 (Order 755 was introduced in October 2011). The full specification also includes fixed effects for CPC classes ($\nu_c$) and year ($\theta_t$). Errors are clustered at the CPC main group level.

The parameter of interest, $\beta_1$, captures how energy storage patent outcomes respond to the policy change that removes barriers for the deployment of energy storage projects. The DiD estimator requires a parallel trends assumption for this estimated effect to have a causal interpretation. $\beta_1$ is identified if, in the absence of treatment, patent outcome of energy storage technologies would have trended in the same way as the patent outcomes of non-storage technologies.

The results from estimating equation 3.3 are shown in Table 3.3. The introduction of Order 755 led to higher patenting activity in energy storage relative to closely related technologies, an increase of nearly 15 patents per year on energy storage technologies using the preferred column (3) estimates which include year and CPC fixed effects. This represents 16% of the average of 92 patents per year in each energy storage CPC main group during the sample period of 2006-2018.

How economically relevant is this estimated effect from a policy perspective? For a comparison with literature in a very different health policy context, Finkelstein (2004) finds that public health policies aimed at increasing vaccination rates led to a 2.5-fold increase in new vaccine clinical trials for vaccine developments that became more profitable as a result of the policy, and their causal estimate accounts for nearly one-third of all new vaccine clinical trials during the study period. In terms of the relationship between market size and innovation, my estimates are smaller in magnitude than the findings of Acemoglu and Linn (2004), again in the context of the pharmaceutical industry. Their estimates suggest that a 1 percent increase in the potential market size for a pharmaceutical drug leads to a 6 percent increase in the total number of new drugs entering the U.S. market. To make the comparison possible, I rely on the findings of Tabari and Shaffer (2020), who show that Order 755 led to an increase in the average size of storage projects, in terms of installed capacity, of 260% to 390% in the FERC regions. A back-of-the-envelope calculation implies a 1 percent increase in project size translates into a .04 to .06 percent increase in storage-related patents, rendering my results an order of magnitude smaller than those of Acemoglu and Linn. While my effects appear smaller than these two papers, both Finkelstein and Acemoglu and Linn find weaker effects when measuring innovation with patents.

Table 3.3: Impact of FERC Order 755 on energy storage patents

| | (1) | (2) | (3) |
|---|---|---|---|
| $StorageCPC \times Post$ | 15.994** | 15.908** | 14.956** |
| | (7.078) | (7.085) | (7.139) |
| Year fixed effects | | ✓ | ✓ |
| CPC fixed effects | | | ✓ |
| Observations | 3,795 | 3,795 | 3,790 |
| Number of Clusters | 337 | 337 | 332 |
| R-Square | 0.025 | 0.026 | 0.966 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Estimates shown are based on equation 3.3. Not shown in the regression (because of the fixed effects in the preferred specification), the estimating equation in column (1) also includes dummies for $StorageCPC_c$ and $Post_t$, and column (2) includes a $StorageCPC_c$ dummy. The sample comprises all patents with at least one US inventor in either energy storage CPC main groups or the closely related technologies, as defined in the text, between 2006 and 2018. Errors are corrected for clustering at the CPC main group level.

One possible concern is that patent outcomes of energy storage technologies were already trending differently relative to the closely related non-storage technologies before the introduction of Order 755. I assess this possibility by estimating a dynamic version of the DiD model replacing the $Post_t$ indicator with year dummy variables:

$$PC_{ct} = \sum_{\tau} \beta_{\tau}(StorageCPC_c \times D_{\tau}) + \nu_c + \theta_t + \varepsilon_{ct}, \tag{3.4}$$

where $D_{\tau}$ are dummies for years $\tau$, and $\tau$ ranges from 2006 to 2018 with 2011 being omitted as the reference period since the policy change occurs near the end of that year. Estimates for the coefficients $\beta_{\tau}$ of the interaction of year dummies with the energy storage CPC indicator from equation 3.4 are shown in Figure 3.4. We see little evidence that patenting in energy storage technologies was in a different trend relative to close non-storage technologies in the years preceding the policy change. The estimated effect occurs rapidly after the policy change, starting from 2012, and is fairly stable over time. The 95% confidence intervals in years post policy change include the estimated magnitudes seen in Table 3.3.

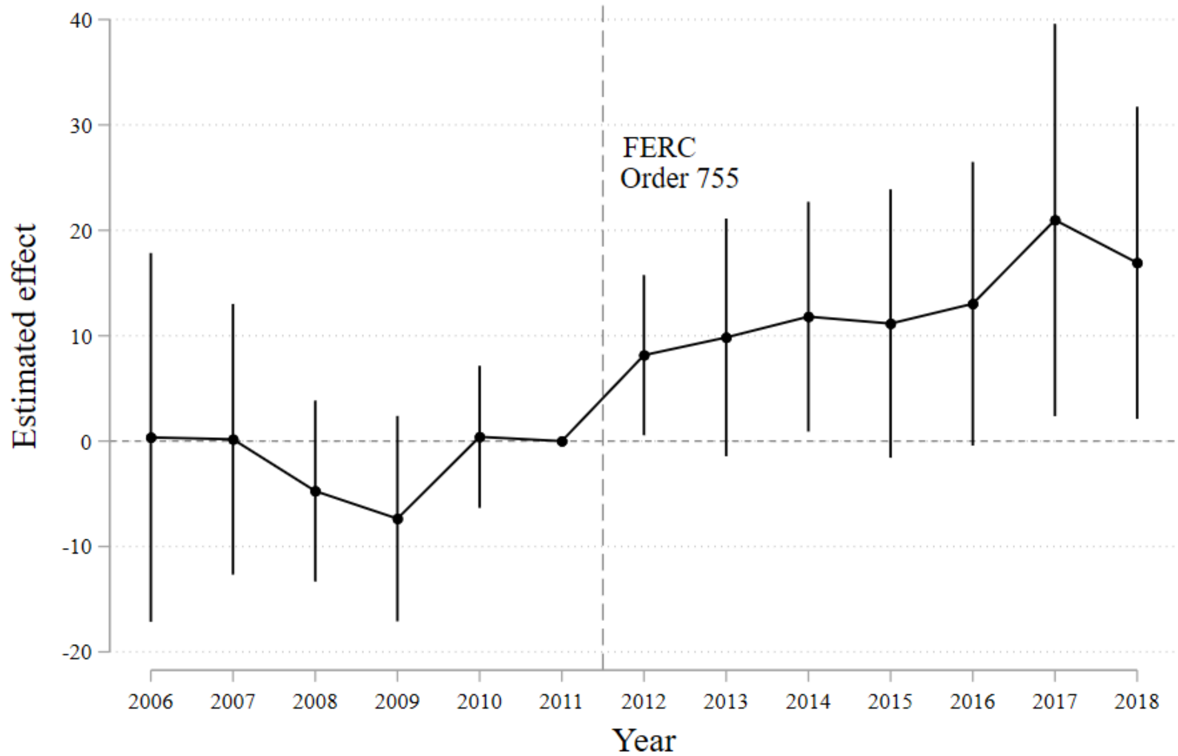



Figure 3.4: Event study: effect of FERC Order 755 on energy storage patents

*Notes:* Estimates shown correspond to the interaction coefficients $\beta_\tau$ from equation 3.4. The reference year is 2011, as the policy change takes place near the end of that year. The sample comprises all patents with at least one US inventor in either energy storage CPC main groups or the closely related technologies, as defined in the text, aggregated annually between 2006 and 2018 to CPC main groups. Errors are corrected for clustering at the CPC main group level.

## 3.5 Conclusion

In this paper, I document a relationship between energy storage innovation and the market size for storage technologies. I subsequently present evidence, from a regulatory change that potentially increased the market size for storage technologies, that energy policy can have a meaningful impact on energy storage innovation as measured by patenting activity.

My results show that the arrival of a new energy storage project causes an immediate but short-lived increase in a county's patenting activity on storage-related technologies. The short-lived nature of the effect suggests that the increase is not the result of local knowledge spillovers since follow-on innovation takes time. I speculate the immediate response is due to early patenting as firms commit to a storage technology prior to a project deployment. This possibility remains open to investigation in future research.

On the policy side, I study the effect on innovation of the introduction of FERC Order 755 in October 2011. Based on this order, frequency regulation resources received a payment for increasing regulation capacity, accuracy, and speed. The policy represented opportunities to grow the market for fast-acting grid-scale energy storage. I employ a difference-in-differences approach and quantify an economically significant increase in the national output of energy storage-related patents attributable to Order 755.

Energy storage resources are gaining importance as intermittent renewable energy sources take up an increasingly bigger role in the supply mix of electrical grids. The message from this paper is that incentivizing the expansion of energy storage projects can indirectly help promote innovation, eventually allowing energy storage to realize its full potential in supporting the transition to clean energy generation.

# Bibliography


Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The environment and directed technical change. *American economic review*, 102(1):131–166.

Acemoglu, D. and Linn, J. (2004). Market size in innovation: theory and evidence from the pharmaceutical industry. *The Quarterly journal of economics*, 119(3):1049–1090.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., and Zhuravskaya, E. (2015). Radio and the rise of the nazis in prewar germany. *The Quarterly Journal of Economics*, 130(4):1885–1939.

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., and Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 124(1):1–51.

Aker, J. C. (2011). Dial "a" for agriculture: a review of information and communication technologies for agricultural extension in developing countries. *Agricultural economics*, 42(6):631–647.

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.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did railroads induce or follow economic growth?: Urbanization and population growth in the american midwest, 1850–1860. *Social Science History*, 34(2):171–197.

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.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of economic perspectives*, 33(3):163–84.

Bridle, L., Magruder, J., McIntosh, C., and Suri, T. (2020). Experimental insights on the constraints to agricultural technology adoption. *Working Paper*.

Broadcasting Publications, Inc. (1935–1959). Broadcasting advertisement yearbook. Accessed online 19-July-2021 from https://worldradiohistory.com/Broadcasting-Yearbook.htm.

Brunner, E. d. (1936). *Radio and the Farmer*. The radio institute of the audible arts.

Burghardt, K., Uhl, J. H., Lerman, K., and Leyk, S. (2022). Road network evolution in the urban and rural united states since 1900. *Computers, Environment and Urban Systems*, 95:101803.

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.

Chan, J. (2022). Farming output, concentration, and market access: Evidence from the 19th-century american railroad expansion. *Journal of Development Economics*, 157:102878.

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.

Chen, C., Restuccia, D., and Santaeulàlia-Llopis, R. (2023). Land misallocation and productivity. *American Economic Journal: Macroeconomics*, 15(2):441–465.

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.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in ghana. *American economic review*, 100(1):35–69.

Coulom, R. (2007). Computing "elo ratings" of move patterns in the game of go. *ICGA journal*, 30(4):198–208.

Coulom, R. (2020). Go Ratings. Accessed online 27-Mar-2020 from https://www.goratings.org/en/.

Crabtree, C. and Kern, H. L. (2018). Using electromagnetic signal propagation models for radio and television broadcasts: An introduction. *Political Analysis*, 26(3):348–355.

Craig, S. (2001). " the farmer's friend": Radio comes to rural america, 1920–1927. *Journal of Radio Studies*, 8(2):330–346.

Craig, S. (2006). " the more they listen, the more they buy" radio and the modernizing of rural america, 1930-1939. *Agricultural history*, pages 1–16.

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.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., and Zhuravskaya, E. (2014). Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics*, 6(3):103–32.

Douglas, G. H. (1987). *The early days of radio broadcasting*. McFarland.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8(4):36–65.

Eckert, F., Gvirtz, A., Liang, J., and Peters, M. (2020). A method to construct geographical crosswalks with an application to us counties since 1790. Technical report, National Bureau of Economic Research.

Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from russia. *American Economic Review*, 101(7):3253–85.

Feng, S. and Lazkano, I. (2022). Innovation trends in electricity storage: What drives global innovation? *Energy Policy*, 167:113084.

Finkelstein, A. (2004). Static and dynamic effects of health policy: Evidence from the vaccine industry. *The Quarterly Journal of Economics*, 119(2):527–564.

Fischer, G., Nachtergaele, F. O., van Velthuizen, H., Chiozza, F., Francheschini, G., Henry, M., Muchoney, D., and Tramberend, S. (2021). Global agro-ecological zones (gaez v4)-model documentation.

Fishback, P. V., Horrace, W. C., and Kantor, S. (2005). Did new deal grant programs stimulate local economies? a study of federal grants and retail sales during the great depression. *The Journal of Economic History*, 65(1):36–71.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, 103(6):1176–1209.

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

Gagliarducci, S., Onorato, M. G., Sobbrio, F., and Tabellini, G. (2020). War of the waves: Radio and resistance during world war ii. *American Economic Journal: Applied Economics*, 12(4):1–38.

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.

Gissey, G. C., Dodds, P. E., and Radcliffe, J. (2018). Market and regulatory barriers to electrical energy storage innovation. *Renewable and Sustainable Energy Reviews*, 82:781–790.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Gowrisankaran, G., Reynolds, S. S., and Samano, M. (2016). Intermittency and the value of renewable energy. *Journal of Political Economy*, 124(4):1187–1234.

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.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522.

Griliches, Z. (1998). Patent statistics as economic indicators: a survey. In *R&D and productivity: the econometric evidence*, pages 287–343. University of Chicago Press.

Gupta, A., Ponticelli, J., and Tesei, A. (2020). Information, technology adoption and productivity: The role of mobile phones in agriculture. Technical report, National Bureau of Economic Research.

Haines, M., Fishback, P., and Rhode, P. (2014). United states agriculture data, 1840-2012. Inter-university Consortium for Political and Social Research [distributor], 2018-08-20. https://doi.org/10.3886/ICPSR35206.v4.

Haines, M. R. (2005). Icpsr 2896 historical, demographic, economic, and social data: The united states, 1790-2000.

Hausman, J. A., Ostro, B. D., and Wise, D. A. (1984). Air pollution and lost work. Working paper series, National Bureau of Economic Research.

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

Holub, F. and Thies, B. (2023). Air quality, high-skilled worker productivity and adaptation: Evidence from github. Mimeo, Available at SSRN: http://ssrn.com/abstract=4398088.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Hutchins, J. (2023). The us farm credit system and agricultural development: Evidence from an early expansion, 1920–1940. *American Journal of Agricultural Economics*, 105(1):3–26.

Igami, M. (2020). Artificial intelligence as structural estimation: Deep blue, bonanza, and alphago. *The Econometrics Journal*, 23(3):S1–S24.

Jia, R. and Ku, H. (2019). Is china's pollution the culprit for the choking of south korea? evidence from the asian dust. *The Economic Journal*, 129(624):3154–3188.

Kantor, S. and Whalley, A. (2019). Research proximity and productivity: long-term evidence from agriculture. *Journal of Political Economy*, 127(2):819–854.

Kim, B., Kim, E., Miller, D. J., and Mahoney, J. T. (2016). The impact of the timing of patents on innovation performance. *Research Policy*, 45(4):914–928.

Krebs, B. and Luechinger, S. (2021). Air pollution, cognitive performance, and the role of task proficiency. *Cognitive Performance, and the Role of Task Proficiency (October 21, 2021)*.

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.

Lai, W., Li, S., Li, Y., and Tian, X. (2022). Air pollution and cognitive functions: evidence from straw burning in china. *American Journal of Agricultural Economics*, 104(1):190–208.

Lanjouw, J. O., Pakes, A., and Putnam, J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. *The journal of industrial economics*, 46(4):405–432.

Le, K. (2020). Land use restrictions, misallocation in agriculture, and aggregate productivity in vietnam. *Journal of Development Economics*, 145:102465.

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.

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.

Manuelli, R. E. and Seshadri, A. (2014). Frictionless technology diffusion: The case of tractors. *American Economic Review*, 104(4):1368–91.

McGowan, D. and Vasilakis, C. (2019). Reap what you sow: Agricultural technology, urbanization and structural change. *Research Policy*, 48(9):103794.

Moretti, E. (2004). Estimating the external return to higher education: Evidence from cross-sectional and longitudinal data. *Journal of Econometrics*, 120(1-2):175–212.

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.

Mueller, S. C., Sandner, P. G., and Welpe, I. M. (2015). Monitoring innovation in electrochemical energy storage technologies: A patent-based approach. *Applied Energy*, 137:537–544.

Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American economic review*, 56(1/2):69–75.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics*, 114(3):941–975.

Nguyen, T., Byrne, R., Trevizan, R., Bastos, A., Wilches-Bernal, F., Concepcion, R., Tonkoski, R., and Headley, A. (2021). Applications and grid services. In *US DOE Energy Storage Handbook*. Sandia National Laboratories.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics*, 94(1):20–36.

Olken, B. A. (2009). Do television and radio destroy social capital? evidence from indonesian villages. *American Economic Journal: Applied Economics*, 1(4):1–33.

Oughton, E., Russell, T., Johnson, J., Yardim, C., and Kusuma, J. (2020). Itmlogic: The irregular terrain model by longley and rice. *Journal of Open Source Software*, 5(51).

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.

Popp, D. (2002). Induced innovation and energy prices. *American economic review*, 92(1):160–180.

PRISM (2011). Prism climate data. *Oregon State University*.

Radio Age (1922–1927). Radio age magazine. Accessed online 19-July-2021 from https://worldradiohistory.com/Radio_Age-20s__Master_Page.htm.

Radio Annual (1938–1950). Stations of the united states. Accessed online 19-July-2021 from https://worldradiohistory.com/Radio_Annual_Master_Page.htm.

Radio Digest (1922–1933). Radio digest magazine. Accessed online 19-July-2021 from https://worldradiohistory.com/Radio_Digest_Master_Page_Guide.htm.

Saavedra, M. and Twinam, T. (2020). A machine learning approach to improving occupational income scores. *Explorations in Economic History*, 75:101304.

Sakti, A., Botterud, A., and O'Sullivan, F. (2018). Review of wholesale markets and regulations for advanced energy storage services in the united states: Current status and path forward. *Energy policy*, 120:569–579.

Schmidheiny, K. and Siegloch, S. (2023). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*.

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.

Squicciarini, M. P. and Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, 130(4):1825–1883.

Strömberg, D. (2004). Radio's impact on public spending. *The Quarterly Journal of Economics*, 119(1):189–221.

Suri, T. and Udry, C. (2022). Agricultural technology in africa. *Journal of Economic Perspectives*, 36(1):33–56.

Tabari, M. and Shaffer, B. (2020). Paying for performance: The role of policy in energy storage deployment. *Energy Economics*, 92:104949.

Tyler, T. F. (1933). *An appraisal of radio broadcasting in the land-grant colleges and state universities*. National committee on education by radio.

USDA (1922–1933). The official record of the united states department of agriculture.

USGS (2017). 1/3rd arc-second digital elevation models (dems)—usgs national map 3dep downloadable data collection.

Wang, T. (2021). Media, pulpit, and populist persuasion: Evidence from father coughlin. *American Economic Review*, 111(9):3064–92.

Wik, R. M. (1981). The radio in rural america during the 1920s. *Agricultural history*, 55(4):339–350.

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.

# Appendix A

## Additional Tables and Figures for Chapter 1

Table A1: Baseline Specification Allowing for Spatial Correlation in Error Term

| Distance cutoff: | 25km | 50km | 100km | 200km |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A: Dependent Variable – Log(farm value/acre)* | | | | |
| $FarmSignal$ | 0.021*** | 0.021*** | 0.021** | 0.021 |
| | (0.000) | (0.001) | (0.030) | (0.109) |
| *Panel B: Dependent Variable – Log(crop value/acre)* | | | | |
| $FarmSignal$ | 0.044*** | 0.044*** | 0.044** | 0.044* |
| | (0.000) | (0.001) | (0.018) | (0.090) |
| Observations (Either Panel) | 13,380 | 13,380 | 13,380 | 13,380 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

*Notes:* Table shows full baseline specification of equation 1.1 with error terms adjusted to allow for spatial correlation following Conley (1999)'s approach with various distance cutoffs.

Table A2: Change in Remoteness

| | (1) Log(farm value/acre) | (2) Log(crop value/acre) |
|---|---|---|
| $FarmSignal$ | 0.014** | 0.052*** |
| | (0.006) | (0.011) |
| $FarmSignal \times NewRoadsDensity$ | -0.002 | -0.006 |
| | (0.003) | (0.006) |
| Full baseline controls and FEs | ✓ | ✓ |
| Observations | 8,826 | 8,826 |
| Number of Clusters | 1,471 | 1,471 |
| Adjusted R-Squared | 0.957 | 0.826 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Table All baseline variables defined as previously in Tables 1.2 to 1.4. $NewRoadsDensity$ measures the number of new roads per km$^2$ of built area between the years 1900 and 1940, standardized with a mean zero and variance one. Standard errors are corrected for clustering at the county level. The reduced number of observations in these regressions relative to the main analysis sample are due to the limited availability of the road density data.

Table A3: Robustness to Alternative Specifications

| Dependent variable: | Log(farm value/acre) | | | Log(crop value/acre) | | |
|---|---|---|---|---|---|---|
| Model: | Baseline | Flexibly control $FarmSignalFree$ | Farm population weighted regression | Baseline | Flexibly control $FarmSignalFree$ | Farm population weighted regression |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $FarmSignal$ | 0.021*** | 0.027*** | 0.026*** | 0.044*** | 0.053*** | 0.042*** |
| | (0.005) | (0.005) | (0.009) | (0.009) | (0.009) | (0.015) |
| Full baseline controls and FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 13,380 | 13,380 | 13,380 | 13,380 | 13,380 | 13,380 |
| Number of Clusters | 2,230 | 2,230 | 2,230 | 2,230 | 2,230 | 2,230 |
| Adjusted R-Squared | 0.958 | 0.958 | 0.963 | 0.829 | 0.829 | 0.835 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Table shows robustness checks on the baseline main results. Columns (1) and (4) reproduce main results previously shown in Table 1.2 and 1.3. Columns (2) and (5) control flexibly for farm signal in free space with a third order polynomial. Columns (3) and (6) weigh the baseline regression with the county's farm population. All regressions include the full set of controls and fixed effects used in the preferred specification.

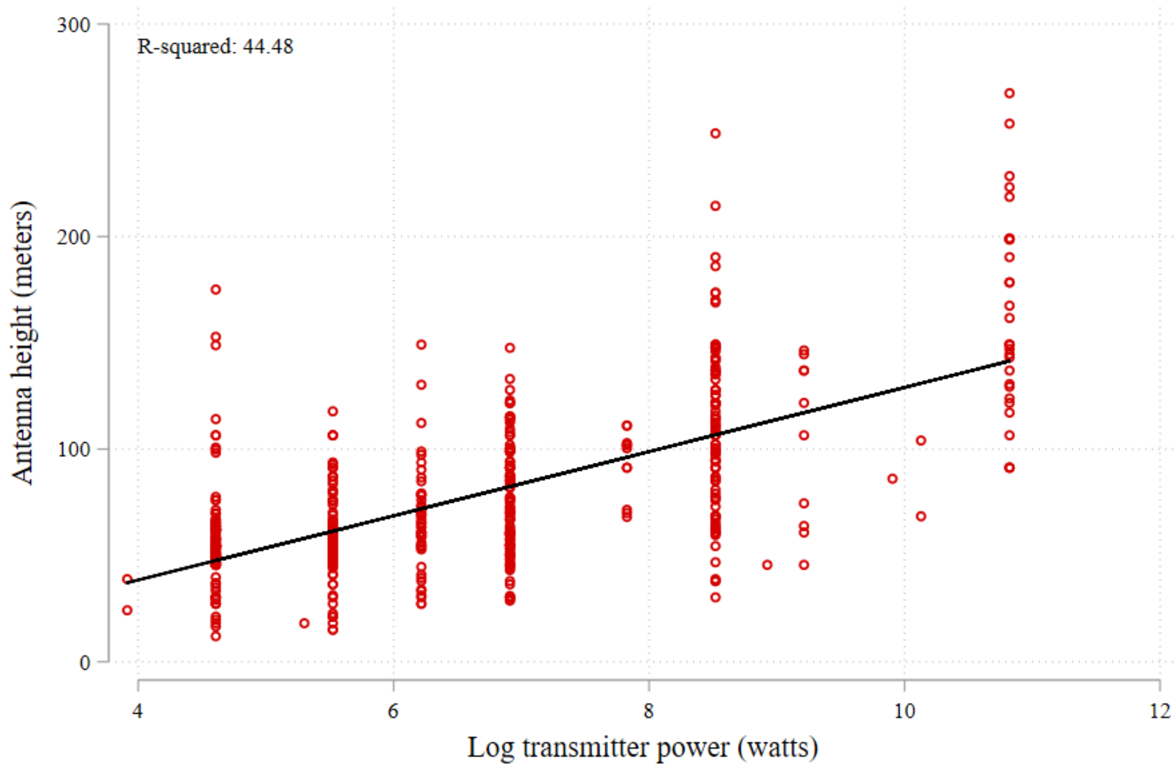



Figure A1: Predicting Missing Antenna Height Values

*Notes:* Antenna height data drawn from the 1940 issue of the *Broadcasting Yearbook* (Broadcasting Publications, Inc., 1959).

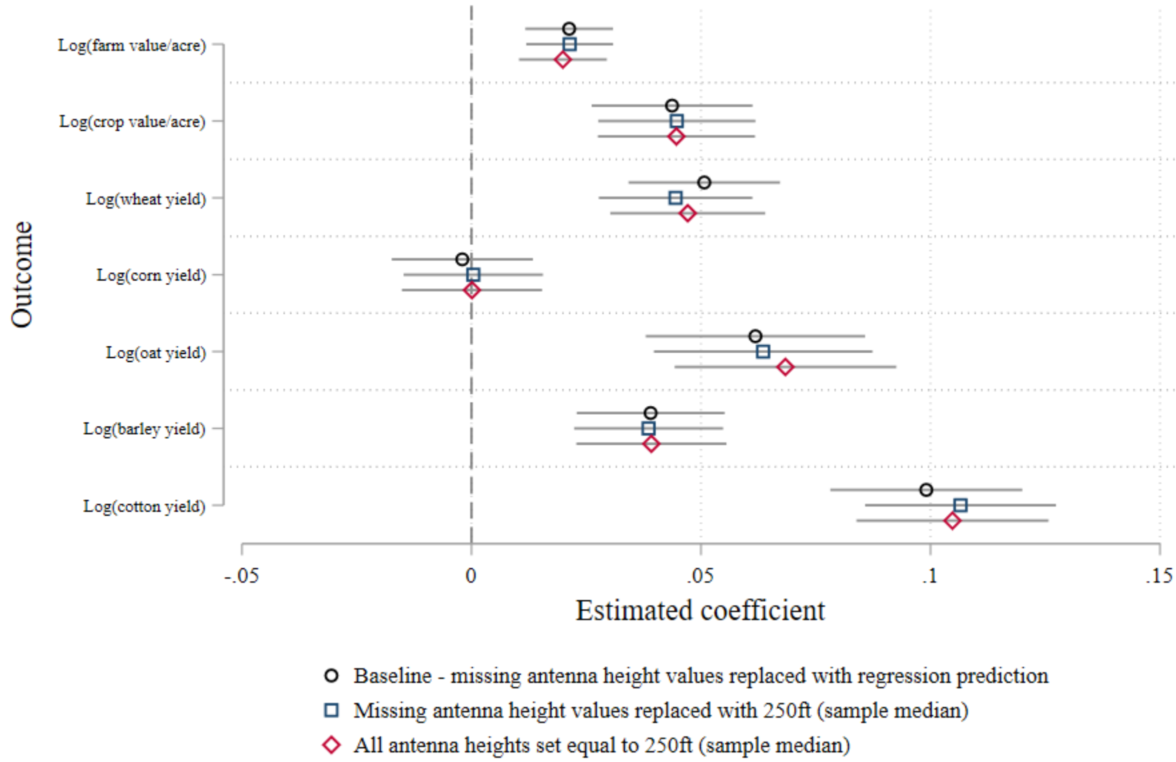



Figure A2: Alternative Treatment of Missing Antenna Height Values

*Notes:* This figure shows the sensitivity of the estimated effect of farm radio to different ways to fill missing values of antenna height. Antenna height data is only available for 1940 and on the baseline signal strength calculations using the ITM and free space propagation models the missing values are replaced with predictions from a simple linear regression of antenna height on log transmitter power ($R^2$ = .44). The figure additionally shows the estimated coefficients when the signal strengths are calculated replacing missing antenna height values with the 1940 median height or when all antenna height values in the sample are replaced with the 1940 median height.

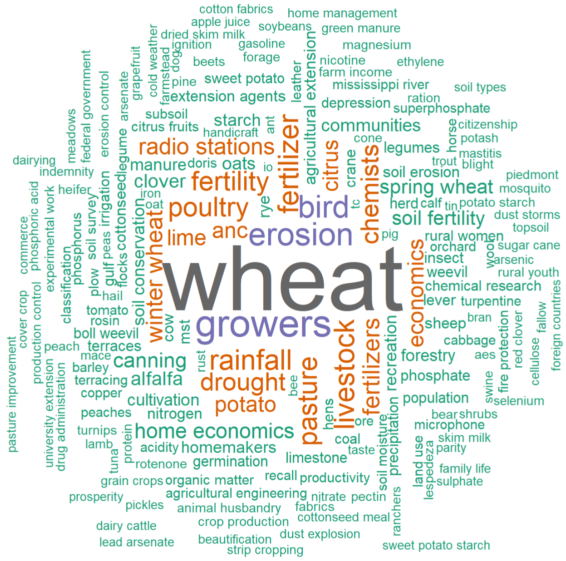
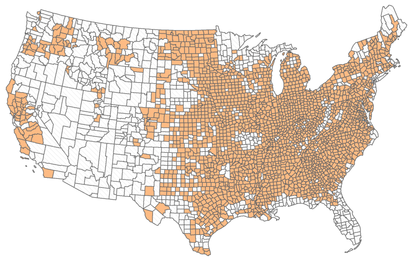


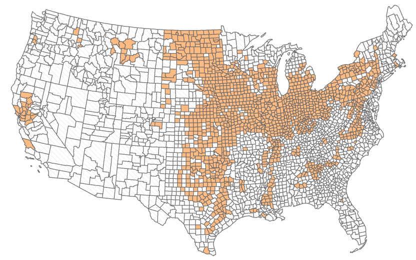

Figure A3: Wordcloud of relevant science terms on transcripts of The National Farm and Home Hour

*Notes:* These transcripts are an extensive (but not comprehensive) database of digitized scripts of the USDA's The National Farm and Home Hour, spanning the years 1929-1942, and available through the USDA's National Agricultural Library. The transcripts are hosted in an Internet Archive collection (url: https://archive.org/details/usda-nationalfarmhomehour), and was accessed in September 18 2021. The word cloud depicts the most commonly found terms – based on the ScienceDirect dictionary of scientific topics – on the set of digitally available scripts.

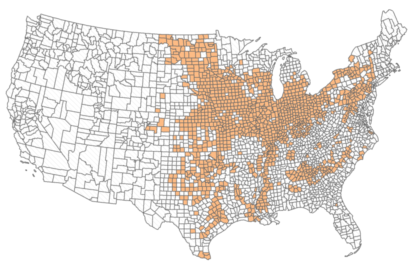
(a) Baseline counties



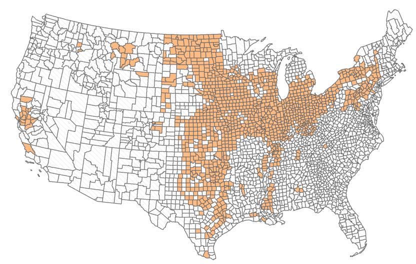
(b) Baseline , high wheat suitability



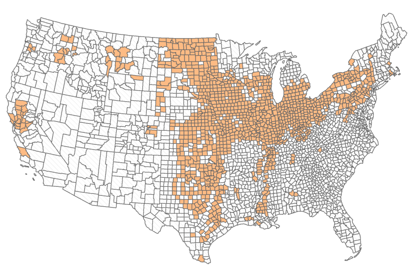
(c) Baseline , high corn suitability



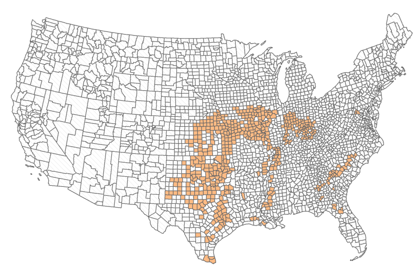
(d) Baseline , high barley suitability



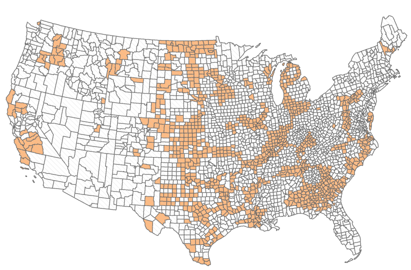
(e) high oat suitability



(f) Baseline , high cotton suitability



(g) Baseline, >100km from farm station



(h) All continental US counties

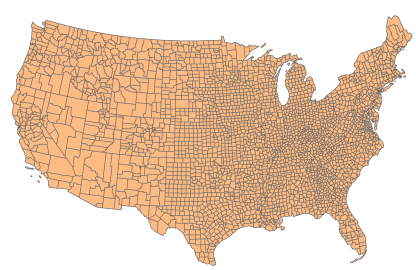


Figure A4: Sample of counties in baseline and robustness checks

*Notes:* Panel (a) baseline corresponds to the main sample described in the Data subsection 3.1. Panels (b) to (f) comprise counties within the baseline with a county averaged suitability index above 50% for the specified crop. Panel (g) comprises counties within the baseline that are further than 100km from the nearest farm radio station in all periods between 1925-1950.

# Exhibit 1. Select Radio Programming Excerpts

**On soil erosion:**

- "In the black land of Texas, some of the greatest cotton lands of the world, we have an erosion experimental farm near the town of Temple. The chief development there last year and the year before was along the line of strips cropping as already mentioned in connection with the Guthrie, Oklahoma, Station work. Under this method which Bennett has explained to you before, farmers plant strips of thick-growing, soil-saving crops, such as oats, sorghum, and sweet clover, along the contours of the field slopes. These are comparatively narrow strips. Then they plant broader strips of the clean-tilled crops, such as cotton and corn, between the strips of soil-saving crops. Practically no erosion or run-off came from the strip-cropped fields at the Temple station."
  – Dr. Henry G. Knight, Chief, Bureau of Chemistry and Soils, for The National Farm and Home Hour, Jan. 11, 1933.

**On weather and crop outlook:**

- "June weather, especially during the latter part of the month, was very trying to man, beast, and many crops over large sections of the country, especially in the States comprising the central valleys and the Northwest. However, a hot, dry spell could hardly have come at a better time, to cause the least amount of damage to staple crops. Winter wheat was largely too far advanced to be seriously harmed, and corn in the principal producing sections had not reached its critical stage of growth. Late spring wheat, oats, other small grains, potatoes, truck, and pastures were less fortunate, especially in the North-Central States, and these suffered considerable damage. Corn was not permanently injured in the main producing sections. In fact, it made exceptional and phenomenal growth, wherever there was sufficient soil moisture and, in general, the crop is in excellent shape at the present time and much ahead of an average season, except in some dry southern sections. In Oklahoma, corn is in a critical stage of growth, and needs moisture badly, while in many other southern localities, especially

in the Southeast, the crop has been damaged by drought. Cotton, while late, continued to make mostly satisfactory growth, but moisture is needed in the northwestern Belt, especially in Oklahoma, and in the Southeast, notably in Georgia and some adjoining sections."
– J.B. Kincer, Meteorologist, Weather Bureau, for The National Farm and Home Hour, July 8, 1931.

- "Taking the country as a whole, the weather was better in August than in July. The result – a 5 percent increase in the crop yield prospects. Although several crops are late and in danger from early frosts or wet weather, an abundant harvest now seems almost assured. The picture isn't equally bright in all sections of the country. Storms along the Louisiana and Carolina coasts caused losses of rice, tobacco, peanuts, and peaches. Dry weather continued through August in an area extending from east central Nebraska to central Colorado, and into late August in central Illinois, Kentucky, and New England, while in the northern and central portions of the Corn Belt and in the Southwest good weather brought marked improvement in the prospects for corn, sorghums, small grains, and other crops. [...] The estimate for September 1 is slightly over 2 and a quarter billion bushels, that's an increase of about 49 million bushels over a month ago [about corn]. The estimate is for nearly 785 million bushels, up more than 20 million bushels in the past month [about wheat]. About 52 million bushels, nearly 3 million less than expected a month ago [about rice]. "
– E. J. "Mike" Rowell, Agricultural Marketing Service, for The National Farm and Home Hour, Sep. 11, 1940.

# Appendix B

## Additional Figures for Chapter 2

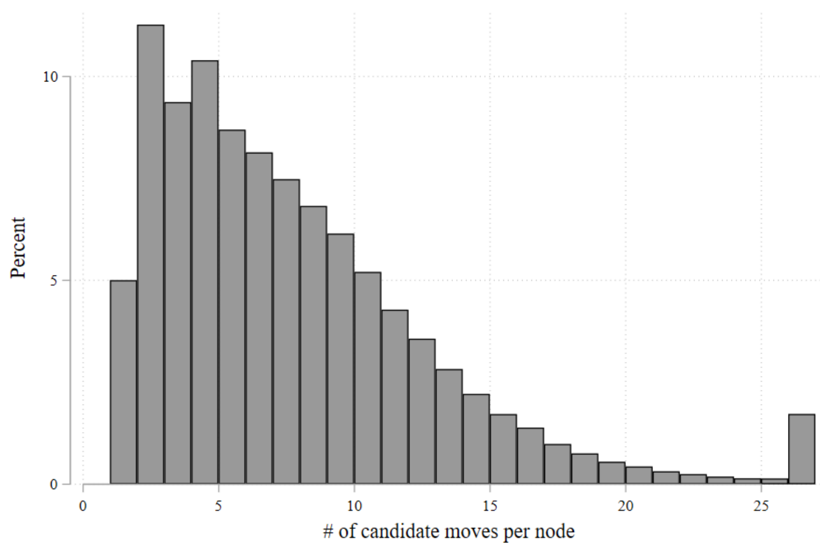



Figure B1: Histogram of candidate moves per node proposed by Leela Zero

*Notes:* 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.

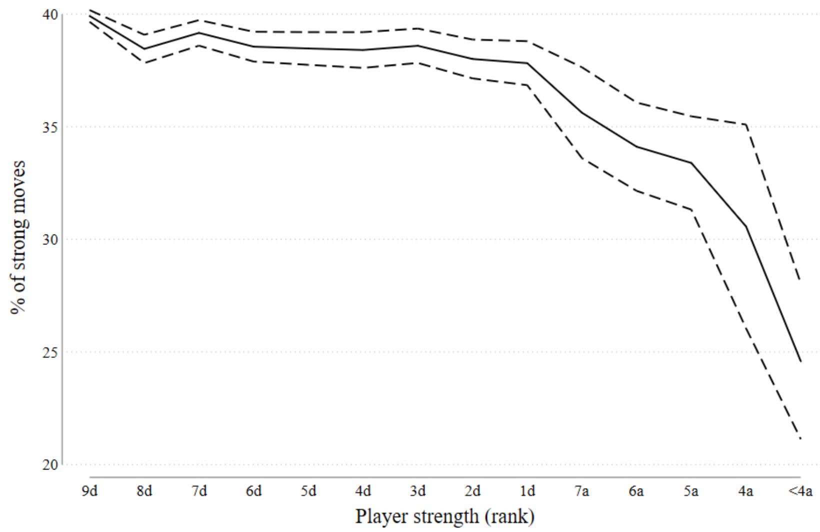



Figure B2: Mean percentage of strong moves per game by player strength

*Notes:* Rank in x-axis decreases in strength from left to right and represents the highest dan degree achieved by a player at game day. Percentage of strong moves is calculated as the percentage of move choices by a player that coincide with the best move suggested by Leela Zero in the mid-game range of moves 100-119 and is averaged at the dan level. Dashed lines plot a 95% confidence interval around the mean, represented by the solid line.

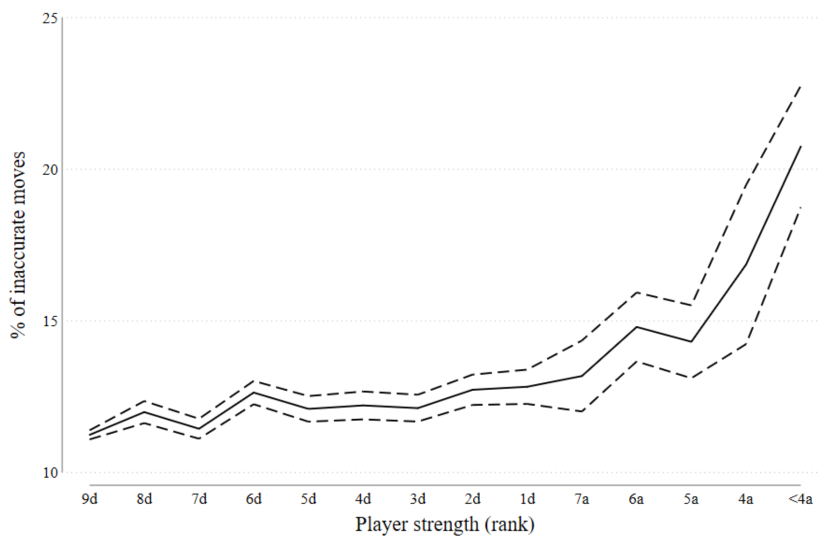



Figure B3: Mean percentage of inaccuracies per game across different ranks

*Notes:* Rank in x-axis decreases in strength from left to right and represents the highest dan degree achieved by a player at game day. Percentage of inaccurate moves is calculated as the percentage of a player's move choices, in the mid-game range of moves 100-119, which are not in the set of candidate moves proposed by Leela Zero, averaged at the dan level. Dashed lines plot a 95% confidence interval around the mean, represented by the solid line.

# Appendix C

## Additional Tables for Chapter 3

Table C1: Effect of new energy storage projects on patenting, using inverse hyperbolic sine

| Dependent Variable = | $IHS$(Energy storage patents) | | | $IHS$(All other patents) | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $NewProjects$ | 0.03705*** | 0.03540*** | 0.02733** | 0.00815* | 0.00220 | 0.00067 |
| | (0.01152) | (0.01123) | (0.01078) | (0.00442) | (0.00468) | (0.00490) |
| County fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | | | ✓ | | |
| Year × ISO/RTO fixed effects | | ✓ | | | ✓ | |
| Year × State fixed effects | | | ✓ | | | ✓ |
| Observations | 33,154 | 33,154 | 33,143 | 33,154 | 33,154 | 33,143 |
| Number of Clusters | 3,014 | 3,014 | 3,013 | 3,014 | 3,014 | 3,013 |
| R-Square | 0.887 | 0.887 | 0.890 | 0.948 | 0.948 | 0.949 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

*Notes:* Estimates shown are based on equation 3.1. Errors are corrected for clustering at the county level.