

NBER WORKING PAPER SERIES

POLLUTION, INFECTIOUS DISEASE, AND MORTALITY:
EVIDENCE FROM THE 1918 SPANISH INFLUENZA PANDEMIC

Karen Clay
Joshua Lewis
Edson Severnini

Working Paper 21635
<http://www.nber.org/papers/w21635>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2015

We thank Martha Bailey, Antonio Bento, Christian Dippel, Walker Hanlon, and seminar participants at the 2015 NBER DAE Summer Meetings, the 2015 CNEH Meetings, the 2015 AERE Summer Conference, the 2015 Annual Meeting of the Population Association of America, McGill, Carnegie Mellon, and USC for valuable comments and suggestions. Edson Severnini would also like to thank the Energy Policy Institute at the University of Chicago (EPIC) for research support and its generous hospitality during his semester-long visit, when part of this project was executed. Karen Clay and Edson Severnini acknowledge financial support from Heinz College and the Berkman Fund at Carnegie Mellon. Joshua Lewis acknowledges financial support from the University of Montreal. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Karen Clay, Joshua Lewis, and Edson Severnini. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic
Karen Clay, Joshua Lewis, and Edson Severnini
NBER Working Paper No. 21635
October 2015, Revised September 2016
JEL No. I15,I18,N32,N52,Q40,Q53

ABSTRACT

This paper uses the 1918-1919 influenza pandemic as a quasi-natural experiment to examine whether air pollution affects susceptibility to infectious disease. The analysis combines the sharp timing of the pandemic with large cross-city differences in baseline pollution measures based on coal-fired electricity generating capacity. The findings suggest that air pollution had a quantitatively important impact on pandemic severity. Had pollution in coal-intensive cities been reduced to the median level, pandemic-related mortality would have been 10 to 14 percent lower, and pandemic-related infant mortality would have been 25 to 40 percent lower. These results have implications for pandemic preparedness and the allocation of scarce resources during an outbreak.

Karen Clay
Heinz College
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
and NBER
kclay@andrew.cmu.edu

Edson Severnini
Carnegie Mellon University
4800 Forbes Ave #2114B
Pittsburgh, PA 15213
ersevernini@gmail.com

Joshua Lewis
University of Montreal
Département de sciences économiques
3150 rue Jean-Brillant
Montréal, QC, H3T 1N8
joshua.lewis@umontreal.ca

1 Introduction

Influenza pandemics are unpredictable but recurring events that create substantial economic disruption and pose a major threat to global health.¹ For example, the 1918-1919 Spanish Influenza Pandemic is estimated to have killed 50 million people worldwide (Crosby, 1989). A pandemic comparable in severity to the Spanish Flu would quickly overwhelm existing healthcare infrastructure.² Thus, preventative strategies and targeted distribution of medical resources are likely to be critical in mitigating the severity of an outbreak.

Recent evidence suggests that air pollution may exacerbate pandemic severity. In randomized control trials, mice exposed to higher air pollution concentrations experienced increased mortality when infected with a common strain of influenza virus (Harrod et al, 2003; Lee et al, 2014). If this relationship holds in human populations, it has important implications for addressing pandemic influenza, given that 87 percent of the world population lives in locations where air pollution exceeds the World Health Organization Guideline.

Despite the importance for policy, identifying how air pollution affects pandemic severity is challenging because of unobserved confounding factors such as prior exposure to infectious disease or vaccination. This paper uses a quasi-natural experiment – the 1918-1919 Spanish Influenza Pandemic – to study the effects of pollution on pandemic mortality. The unanticipated, short-lived, and severe outbreak killed an estimated 675,000 Americans. The H1N1 strain was new, so individuals did not have

¹The World Bank estimates that the global costs of a severe influenza pandemic could exceed \$3 trillion (World Bank, 2013). Given the non-zero probability of a pandemic in any given year, the annual expected costs are estimated to be over 700,000 deaths worldwide, and 0.7 percent of global income (Fan, Jamison, and Summers, 2016).

²In the United States, Meltzer, Cox, and Fukuda (2005) estimate that the next influenza pandemic could lead to 89,000 to 207,000 deaths; 314,000 to 734,000 hospitalizations; 18 to 42 million outpatient visits; and 20 to 47 million additional illnesses.

immunity, and there were no vaccines or effective medical treatments. Moreover, the influenza virus spread rapidly, limiting the scope for viral evolution. As a result, cohorts exposed to the pandemic experienced a common health shock. There were wide differences in air pollution across cities, in part due to differences in the available inputs for electricity generation. Thus, these cohorts were exposed to markedly different levels of air pollution prior to the outbreak of the pandemic.

The research design combines the sharp timing of the pandemic with large cross-city differences in baseline pollution levels to identify the impact of air pollution on pandemic severity. The empirical analysis draws on information on infant and all-age mortality for a panel of 183 American cities, representing more than one-half of the urban population and one-quarter of the total population, for the years 1915 to 1925. These data are linked to newly digitized detailed 1915 data on the location, capacity and type of generation for all electricity plants with at least 5 megawatts of capacity. Our preferred measure of exposure is total coal-fired electricity capacity within 30 miles of a city. It is highly correlated with measures of bituminous coal consumption and can be thought of as a proxy for city-level pollution.

The results show that air pollution exacerbated the severity of the pandemic. Coal-fired capacity is positively and statistically significantly related to both infant and all-age mortality in 1918. These effects are quantitatively important. Differences in coal-fired capacity can account for 35 to 46 percent of the cross-city variation in pandemic infant mortality and 22 to 32 percent of the cross-city variation in all-age mortality. To illustrate the magnitude of the main estimates, we compute the effect of moving all cities above the median in coal-fired generating capacity to the median. In this scenario, pandemic-related infant mortality would have been 25 to 40 percent lower, and 3,283 to 5,232 infant deaths would have been averted. Meanwhile, all-age mortality would have been 10 to 14 percent lower, and 15,617 to 22,314 all-age deaths

would have been averted.

We examine the sensitivity of the main findings to a number of potential threats to identification. A first concern is that less healthy populations sorted into heavily polluted and also experienced differential increases in mortality in 1918. We include a variety of measures of city-level health interacted with 1918 to allow for differences in pandemic severity according to baseline population health. We also explore heterogeneity in the effects of pollution on pandemic mortality across cities with different baseline levels of health capital. The results show that differences in baseline population health cannot explain the relationship between coal capacity and pandemic severity. A second threat to identification is that the timing of pandemic onset may have been spuriously correlated with local pollution levels. If the timing of onset affected pandemic-related mortality, either because of viral evolution or the local response to the outbreak, this correlation could bias the baseline estimates. We interact the 1918 fixed effect with a number of determinants of pandemic timing, including city proximity to World War I bases, the local railway network, and the week of initial onset. The findings are not sensitive to these controls.

A third concern is that mortality in high pollution cities was particularly elevated in 1918 for reasons unrelated to the pandemic, such as increased wartime production. To examine this issue, we estimate the impact of coal-fired capacity on excess mortality in non-pandemic years.³ We find no relationship between coal-fired capacity and mortality in non-pandemic years. Notably, the 1917 interaction effect is small and statistically insignificant, despite the fact that World War I mobilization was under way. Coal-fired capacity also had a positive impact on mortality in 1919, although the estimate is smaller in magnitude. These results are consistent with the timing

³These models control for the direct impact of air pollution on mortality with city fixed effects and a range of time-varying city-level controls.

of the pandemic, which was most severe in the fall of 1918, but continued to affect mortality through the first quarter of 1919. There is also no impact of hydroelectric capacity on pandemic-related mortality. Hydroelectric capacity serves as a placebo test, because it generates electricity but is emission free.

By studying the relationship between air pollution and pandemic-related mortality, this paper contributes to the literature on determinants of the spread and severity of influenza pandemics. The 1918-1919 Influenza Pandemic continues to be widely studied because of its relevance for the prevention of future outbreaks.⁴ The risks posed by a severe influenza pandemic are substantial and are unlikely to be met by existing medical infrastructure. The results of this paper thus have important implications for pandemic preparedness and the allocation of scarce resources during an outbreak. Mitigation strategies could include both air pollution abatement, and the distribution of vaccines, antiviral drugs, and medical supplies to high pollution locations.

This paper also contributes to the literature on air pollution and mortality by providing some of the first evidence on the health interaction between air pollution and infectious disease. A number of studies have shown a causal link between air pollution and mortality (Chay and Greenstone, 2003a, 2003b; Currie and Neidell, 2005). Because these studies typically rely on short-term variation in air pollution to identify the health impact, they are unable to identify the extent to which air pollution increases susceptibility to other infrequent negative health shocks. Our findings show that baseline exposure to air pollution substantially exacerbated the mortality effects of the 1918 Influenza Pandemic. From a policy perspective, the presence of these

⁴A large medical literature has sought to understand the particular characteristics of the H1N1 strain responsible for the pandemic (see Taubenberger and Morens, 2006, for a discussion). Beginning with Almond (2006), economists have also used the sharp timing of the pandemic to examine the long-term outcomes of survivors (see Nelson, 2010; Neelsen and Stratmann, 2012; and Lin and Liu, 2014).

interaction effects suggests that there may be considerable co-benefits of pollution abatement that are not accounted for by conventional evaluations.

The paper proceeds as follows. Section 2 provides background on the link between air pollution and pandemic-severity and discusses the history of the 1918 influenza pandemic and air pollution in early 20th century United States. Section 3 describes the data. Section 4 introduces the empirical strategy. Section 5 presents the main findings. Section 6 reports a variety of robustness exercises; and section 7 concludes.

2 Background

2.1 Air Pollution, Health Capital, and the Pandemic

In the Appendix, we develop a simple model of health capital to examine the channels through which exposure to air pollution affected the severity of the 1918-1919 Influenza Pandemic. The model provides two main insights. First, differences in population pollution exposure may cause a common health shock, such as the pandemic, to have heterogeneous effects on mortality. Second, both contemporaneous and pre-pandemic pollution exposure can influence pandemic severity. As discussed below, it is likely that contemporaneous pollution exposure weakened individual immunity and increased susceptibility to the pandemic. Meanwhile, baseline pollution exposure could exacerbate or mitigate the impact of the pandemic. On the one hand, pre-pandemic pollution exposure will tend to lower health capital, leaving individuals more vulnerable to the consequences of a negative health shock. On the other hand, if pollution increased pre-pandemic mortality among individuals with marginal health, the average health stock in the local population could increase. For adults, both channels can influence pandemic-related mortality, since that the stock of health

capital will depend on lifetime pollution exposure. For infants, contemporaneous pollution exposure is likely to be more important, since infant health capital depends on pre-pandemic pollution indirectly through the mother’s lifetime exposure.

There is a large literature that examines the direct impact of air pollution on infant mortality and all-age mortality (Chay and Greenstone, 2003a, 2003b; Currie and Neidell, 2005; Pope et al, 2004; DelFino et al, 2005). These studies typically identify the effect of contemporaneous pollution exposure on health, but generally have not examined the health impact of long-term pollution exposure or the extent to which pollution interacts with other negative health shocks.⁵ For infants, prenatal exposure affects health through the timing of birth and birthweight (Currie and Walker, 2011), and postnatal exposure has been linked to Sudden Infant Death Syndrome and to respiratory illness (Woodruff et al, 2008; Arceo-Gomez et al, 2012). Air pollution affects adult mortality primarily through cardiovascular disease and respiratory disease (Hoek et al, 2013).⁶

There is emerging evidence that suggests that air pollution may exacerbate pandemic severity. In randomized control trials, mice exposed to higher levels of particulate matter (PM) experienced increased mortality when infected with a common strain of the influenza virus (Hahon et al, 1985; Harrod et al, 2003, Lee et al, 2014). Microbiology studies of respiratory cells also identify a link between pollution exposure and respiratory infection. Respiratory epithelial cells are the primary site

⁵A small number of recent studies have linked childhood exposure to pollution and long-term outcomes (see Currie et al, 2014, for an overview).

⁶In their review and American Heart Association statement, Brook et al (2010) find: “Most, but not all, epidemiological studies corroborate the elevated risk for cardiovascular events associated with exposure to fine PM <2.5 μm in aerodynamic diameter (PM_{2.5}). PM_{2.5} generally has been associated with increased risks of myocardial infarction (MI), stroke, arrhythmia, and heart failure exacerbation within hours to days of exposure in susceptible individuals.” The mechanism are still being explored, but they summarized the evidence up to that point: “Air pollutants have been linked with endothelial dysfunction and vasoconstriction, increased blood pressure (BP), prothrombotic and coagulant changes, systemic inflammatory and oxidative stress responses, autonomic imbalance and arrhythmias, and the progression of atherosclerosis.”

for influenza virus infection and replication, and PM exposure increases the viral-load post-infection (Jaspers et al, 2005). In experimental models, PM has also been shown to reduce the host response to bacterial infections of the lungs through altered bacterial clearance (Jakab, 1993), an effect that may have been particularly severe during the 1918 epidemic, when mortality was often caused by a secondary infection, such as bacterial pneumonia.⁷

2.2 Air Pollution in Early 20th Century American Cities

Historical evidence suggests that air pollution was a problem and that there were substantial differences in air quality across cities. As smoke became significant, cities often passed legislation aimed at reducing it. In 1912, the Bureau of Mines reported that 23 of 28 cities with populations over 200,000 were trying to combat smoke (Goklany, 1999, p. 15). The top panel of Table A.1 lists the 23 cities with smoke problems and the 5 cities that did not have smoke problems, because they used relatively little coal. The bottom panel of Table A.1 reports cities that adopted legislation to combat smoke problems. Dozens of smaller cities also passed legislation.

Newspapers frequently discussed the ‘smoke nuisance’. Although systematic cross-city information on pollution levels was not available until the mid-1950s, intermittent monitor readings during the early 20th century suggest the problem was severe and varied widely across cities. The top panel of Table A.2 reports TSP concentrations across various cities in the early 20th century. TSP concentrations in these cities were similar to levels in Chinese cities from 1980-1993. These levels are six times higher than the average annual TSP threshold and twice the maximum daily threshold

⁷Contemporary researchers noted the devastation to the lungs of influenza victims. At a discussion reported in the *Journal of the American Medical Association*, pathologists noted that “the lung lesions, complex and variable, struck one as being quite different in character to anything one had met with at all commonly in thousands of autopsies one had performed during the last 20 years” (Ireland, 1928, p.150).

initially set under the Clean Air Act Amendments of 1970.⁸ There was considerable variation in air quality across cities. For example, TSP concentrations in Chicago were twice the levels found in Detroit.

Coal-fired capacity is a useful proxy for city-level air quality. Coal consumption for electricity generation was an important contributor to urban air pollution. Cities near large bituminous coal deposits were heavily dependent on coal-fired capacity, while cities in areas where topographical conditions favored its use were heavily dependent on hydroelectric capacity. The majority of power plants were in urban areas, and emissions were dispersed locally through stacks that were below 75 meters in height on average (Hales, 1976, Figure 4, p.10).⁹ Figure A.2 reports the relationship between total state level coal consumption and coal-fired and hydroelectric capacity in 1917. There is a strong positive relationship between coal-fired capacity and total coal consumption at the state-level. There is no link between hydroelectric capacity and overall coal use.

2.3 The 1918-1919 Influenza Pandemic

The influenza pandemic of 1918-1919 was brief, but severe. Estimates of world-wide fatalities range from 50-100 million (Crosby, 1989; Johnson and Mueller, 2002). In the United States, fatalities were between 675,000 and 850,000. Figure 1 reports national influenza and pneumonia deaths by month for the 1918-1919 period and the corresponding months for the previous 5 years. Pandemic-related mortality was particularly elevated between October 1918 and January 1919. This four-month period accounted for over 90 percent of pandemic-related deaths.

⁸Under the 1970 CAAA, the EPA designated a county as nonattainment if TSP concentrations exceeded either 1) an annual geometric mean concentration of $75 \mu\text{g}/\text{m}^3$, or 2) the second highest daily concentration exceeded $260 \mu\text{g}/\text{m}^3$.

⁹Electricity generation accounted for 20 percent of total domestic coal consumption in 1920 (Historical Statistics, 1976, p.591, 824).

The pandemic was caused by the H1N1 virus. Unlike the seasonal flu, which is typically caused by slight variations in pre-existing strains, the H1N1 virus had not previously been introduced to the human population. As a result, all individuals lacked immunity to the virus. Approximately 30 percent of the U.S. population contracted the H1N1 virus in 1918-1919 (Collins, 1931). Fatality rates for those who contracted the virus were greater than 2.5 percent, which is far higher than usual (0.1 percent). The Spanish Flu was also characterized by an unusual ‘W’ age distribution of mortality (see Figure A.1).

The pandemic spread rapidly throughout the country. The most serious wave originated in Camp Devens near Boston in the first week of September 1918. Figure 2 documents the timing of pandemic onset, based on information compiled across 376 localities by Sydenstricker (1918). The pandemic had surfaced in most East Coast cities by mid-September, and then moved westward, diffusing nationwide by early October. The rapid spread meant that there was limited evolution of the virus.

Medical and public health interventions seemed to have been ineffective. Antibiotics had not yet been developed and so could not be used to treat the bacterial pneumonia that often developed. Medicine had little to offer beyond palliative care, and hospitals were quickly overwhelmed. Municipalities were often slow to adopt preventative measures that included bans of public gatherings, regulations against spitting in public, and encouragement to wear masks. There is debate about the impact of public interventions on the spread of the pandemic. Most researchers consider the public response to have been inadequate to have meaningfully reduced transmission rates (Brainerd and Siegler, 2003; Crosby, 1989). On the other hand, Bootsma and Ferguson (2007), find that public health interventions reduced pandemic mortality across 43 U.S. cities. In the empirical analysis, we exploit information on the timing of the pandemic onset to explore whether the delayed public response contributed to

pandemic severity.

There were wide cross-state and cross-city differences in pandemic severity. For example, mortality was more than twice as high in Pennsylvania relative to neighboring Ohio. Mortality in Dayton, Ohio was 80 percent higher than in Columbus, Ohio (Huntington 1923, table 7). Although researchers have commented on the differences, there is little consensus on the underlying causes (Huntington 1923, Crosby 1989, Kolata 1999, Brainerd and Siegler 2003).

3 Data

Data on infant and all-age mortality are combined with newly digitized information on electricity generation and additional census data on city characteristics. Infant and all-age mortality rates were digitized for a panel of 183 American cities for the period 1915-1925.¹⁰ These data comes from the *Mortality Statistics*, and covers over one quarter of the U.S. population and over half of the urban population.

These health outcomes are linked to a measure of city-level pollution. Data from a 1915 federal report on the location and capacity of coal-fired and hydroelectric power stations with installed capacity of at least 5 megawatts were digitized (U.S. Department of Agriculture, 1916). Using GIS software, these data were combined with information on city locations to construct a measure of local exposure to pollution from coal-fired electricity generation. Our preferred measures are total coal-fired

¹⁰Price Fishback generously provided these data. We begin with an initial sample of 286 cities with a population of at least 20,000 in 1921. From this sample, we drop 88 cities with missing information on covariates, and exclude an additional 15 cities located in states that did not use coal for electricity for generation, leaving us with a final sample of 183 cities. The results are not sensitive to the exclusion of cities with missing covariates. Infant mortality is defined as the number of infant deaths per 1,000 live births, all-age mortality is the number of total deaths (including infant deaths) per 10,000 population. Because we lack annual city-level information on births and population, the rates are calculated by dividing annual deaths by total births and population in 1921. Similar unreported results were found when mortality rates were constructed using 1910 county-level population.

capacity and hydroelectric capacity within 30 miles of each city-centroid. This distance was chosen to capture the fact that the majority of power plant emissions are dispersed locally.¹¹ The sensitivity of the results to this particular cutoff is explored in the empirical analysis. Figure A.3 displays the sample of cities and the location of coal-fired and hydroelectric power plants. The geographic pattern in the use of coal and hydropower reflects the availability of inputs. Coal-fired power was concentrated in the Midwestern states with abundant coal resources. Pennsylvania, Ohio, and Indiana were almost entirely reliant on coal for electricity. Hydroelectricity was prevalent where topographical conditions favored its use.

These data are linked to county-level demographic and economic characteristics drawn from the census of population, and census of manufacturing (Haines and ICPSR, 2010). Demographic controls include county population density, urban population, and the share of white residents in 1910.¹² Economic covariates include employment in manufacturing in 1910, and manufacturing payroll per worker in 1900.¹³ Additional controls include county-centroid longitude and latitude.

4 Empirical Strategy

The empirical analysis evaluates whether variation in mortality during the influenza pandemic was related to local levels of pollution. We estimate the following

¹¹For a review of the mechanics of airborne pollutant transport, see Seinfeld and Pandis (2012). Recent evidence from Illinois found that over 40 percent of PM_{2.5} exposure occurs within 30 miles of a power plant (Levy et al., 2002). Historically, air pollution would have been substantially more localized, given the increase in power plant smoke-stack heights that has occurred over the past 50 years. Hales (1976, Figure 4, p.10).

¹²Similar results are found whether population density is measured using county population in 1910 or city population in 1921.

¹³Data on manufacturing payroll is not available in 1910.

model:

$$\begin{aligned}
\text{Log}(\text{MortRate}_{ct}) &= \alpha + \beta_{\text{Coal}}(I(\text{Year}_t = 1918) \times \text{Log}(\text{CoalCap30mile}_c)) \\
&+ \beta_{\text{Hydro}}(I(\text{Year}_t = 1918) \times \text{Log}(\text{HydroCap30mile}_c)) \quad (1) \\
&+ \lambda_t + \lambda_t Z_c + \lambda_t X_c + \lambda_t W_c + \eta_c + \phi_s t + \epsilon_{ct}
\end{aligned}$$

where the dependent variable, $\text{Log}(\text{MortRate}_{ct})$, denotes the logarithm of the infant or all-age mortality rate in city c in year t . The term η_c denotes a vector of city fixed effects, λ_t denotes a set of year fixed effects, and $\phi_s t$ is a linear state trend which allows for differential trends in mortality across states. Standard errors are clustered at the city-level to adjust for heteroskedasticity and within-city autocorrelation.

The term $I(\text{Year}_t = 1918)$ is a dummy variable for the year 1918, which is meant to proxy pandemic exposure. The term $\text{Log}(\text{CoalCap30mile}_c)$ denotes the logarithm of coal-fired capacity within 30 miles of the city-centroid, our measure of local exposure to air pollution.¹⁴ The interaction term of interest, $I(\text{Year}_t = 1918) \times \text{Log}(\text{CoalCap30mile}_c)$, identifies the differential impact of the pandemic on mortality in cities differing in levels of coal-fired capacity. This interaction term takes the form of a Bartik-style estimate, in which a time-varying national health shock affects locations differentially according to initial local differences in air pollution. A positive estimate of β_{Coal} would suggest that exposure to power plant emissions exacerbated the impact of the influenza outbreak. Meanwhile, the estimate β_{Hydro} captures the impact of emissions-free capacity on pandemic severity. In some specifications, we also include the full vector of coal-fired and hydroelectric capacity-year fixed effect interaction terms. The coal capacity interaction effect in 1919 is of particu-

¹⁴The log-log specification is motivated by recent research that documents a concave (“supra-linear”) concentration-response relationship between pollution and mortality (Goodkind, Cogglin, and Marshall, 2014; Pope, Cropper, Coggins, and Cohen, 2015). Alternative functional forms are explored in the robustness checks.

lar interest, since 17 percent of pandemic-related mortality occurred between January and April, 1919.

The identifying assumption is that, conditional on covariates, unobservable determinants of mortality are not correlated with $I(Year_t = 1918) \times \text{Log}(CoalCap30mile_c)$. In practice, this condition requires that excess mortality in 1918 did not vary systematically across high and low coal cities for reasons unrelated air pollution. There are three main threats to this identification strategy, which we address in turn.

First, city-level pollution may have been systematically related to characteristics of the local population that influenced pandemic severity. Cross-city differences in population characteristics could bias the results if i) coal-fired capacity were correlated with these local characteristics, for example, if less healthy individuals sorted into heavily polluted cities, and ii) these characteristics directly influenced the severity of the pandemic, for example, if individuals with marginal health experienced a differential rise in mortality in 1918. To address this concern, we interact year fixed effects with baseline demographic and economic covariates, and a range of city-level health conditions, allowing for differences in pandemic mortality according to these baseline characteristics.

Second, the timing of pandemic onset may have influenced pandemic severity, either because of evolution in the virus or the public response to the outbreak. Some accounts suggest that the virulence of the H1N1 strain weakened over the course of several weeks (Crosby, 1989). As a result, cities that experienced later outbreaks may have been exposed to a weaker virus. In addition, cities that experienced later outbreaks may have been better able to respond to the pandemic with preventative policies to mitigate disease transmission. If the timing of pandemic onset was spuriously correlated with city-level pollution, evolution in pandemic severity could lead

the estimates in equation (1) to be biased. To address this concern, we interact the 1918 fixed effect with factors related to pandemic timing including city proximity to World War I bases, the local railway network, and the week of pandemic onset.

A final concern is that mortality in high pollution cities was particularly elevated in 1918 for reasons unrelated to the pandemic, such as increased wartime production. Although, nationwide coal consumption rose just slightly in 1918, it is possible that industrial wartime production (and pollution) was directed to coal-intensive cities, which may have directly contributed to mortality in 1918, independently of the pandemic. To address this concern, we examine the impact of coal-fired capacity on mortality in non-pandemic years. We also estimate models that include the interaction effect $I(Year_t = 1918) \times \text{Log}(HydroCap30mile_c)$, which allows for differences in pandemic mortality according to local hydroelectric capacity. Hydroelectric capacity serves as a placebo test, because it generates electricity but is emission free.¹⁵

5 Results

5.1 Coal-Fired Capacity and Pandemic-Related Mortality

Before turning to the main analysis, Figure 3 presents the cross-city distribution of residual mortality by year.¹⁶ The figure captures the extent to which mortality deviated from its predicted trend in each year between 1915 and 1925. The effect of the pandemic is evident in the figure. In 1918, infant mortality exceeded its trend by 19 percent and all-age mortality exceeded its trend by 35 percent. The flattening of

¹⁵A caveat to this comparison is the common support, since many cities relied exclusively on either coal or hydro electricity. If differences across these two groups of cities influenced the severity of the pandemic, a comparison of estimates for coal and hydroelectric capacity will not necessarily capture the effect of air pollution on pandemic mortality.

¹⁶For each city-year observation, residual mortality is calculate as the difference between observed mortality and mortality predicted based on a linear city-specific trend.

the distribution in 1918 reflects the wide cross-city differences in pandemic-severity. Figures 4 plots residual infant and all-age mortality in 1918 for cities above and below median coal-fired capacity. Pandemic-related mortality rates were 6 to 7 percentage points higher in above-median coal capacity cities.

In Table 2, we examine the extent to which mortality in 1918 was related to local coal-fired and hydroelectric capacity. The top panel presents results for infant mortality and the bottom panel presents results for all-age mortality. Column (1) includes city and year fixed effects and a linear state trend.

To address the concern that city-level pollution was systematically related to local characteristics that influenced pandemic severity, column (2) includes baseline demographic factors (population density, percent urban, and the fraction white in 1910) interacted with year fixed effects. These covariates capture differences in pandemic severity according to local population characteristics. Column (2) also include controls for city longitude and latitude interacted with year fixed effects, to address the second identification issue that the geographic spread of the virus was spuriously correlated with its severity.

In column (3), we further address the issue that city-level pollution was related to local determinants of pandemic severity. We include covariates for baseline manufacturing employment and wages interacted with year fixed effects, to allow the pandemic to differ according to local manufacturing activity. These models also control for annual city-level expenditure on health and sanitation.¹⁷

Column (4) addresses the third identification issue that pollution was related to excess mortality in 1918 for reasons unrelated to the pandemic. We include the full vector of interaction effects between coal-fired and hydroelectric capacity and

¹⁷Given that city-level health expenditure in 1918 could have been influenced by pandemic-severity, we estimate regressions with and without these covariates.

year fixed effects. These regressions allow us to explore whether 1918 mortality was particularly elevated in high coal cities because of wartime mobilization, and to assess the common trends assumption.

In the top and bottom panels, across all four specifications, the coefficients on $I(\text{Year}_t = 1918) \times \text{Log}(\text{CoalCap30mile}_c)$ are positive and highly significant. For infant mortality, the coefficients range from 0.0286 to 0.0349. For all-age mortality, the coefficients range from 0.0216 to 0.0233. Meanwhile, there is no significant relationship between hydroelectric capacity and mortality in 1918. In all but one specification, the estimates of β_{Coal} and β_{Hydro} are statistically different from each other. Together, these results indicate that coal-fired capacity captures exposure to air pollution and not some other characteristic related to electricity production.¹⁸

Figures 5 and 6 report the full set of coal-year interaction effects from the generalized version of equation (1) reported in column 4 of Table 2. These models allow coal-fired capacity to affect excess mortality in all non-pandemic years (the corresponding estimates are reported in Table A.3), and provide an opportunity to judge the common trends assumption underlying the research design. There is a strong effect of coal-fired capacity on mortality in 1918, and to a lesser extent in 1919, consistent with the timing of the pandemic. All but one of the non-pandemic year interaction effects are statistically insignificant, and there is little evidence of differential trends in mortality by coal-fired capacity. Importantly, there is no evidence of an interaction effect in 1917. Since mobilization for World War I was under way in 1917, this result provides confidence that the main results were not driven by an expansion in wartime air pollution in high coal cities.

The estimated effects are large, and imply that poor air quality significantly exacerbated the severity of the pandemic. Multiplying the point estimates by the cross-city

¹⁸In Section 6, we further explore the robustness of these findings.

standard deviation in coal-fired capacity, we calculate that differences in coal-fired capacity can account for 37 to 46 percent of the cross-city variation in pandemic-related infant mortality, and 33 to 35 percent of the cross-city variation in pandemic-related all-age mortality.

5.2 Pollution, Health Capital, and Pandemic Severity

In this section, we examine whether differences in city-level health capital can account for the effects reported in Table 2. Differences in city-level health conditions could bias the results if they were related to both local coal-fired capacity and excess pandemic mortality. As we demonstrate in the Appendix, the direction of this potential bias is ambiguous. High pre-pandemic mortality rates could reflect lower baseline health capital, leaving the population particularly vulnerable to the effects of a negative health shock. On the other hand, high pre-pandemic mortality rates could raise average city-level health capital by culling the least healthy members of the population.¹⁹

To assess the robustness of the main estimates to differences in baseline health capital, we interact measures of baseline city-level health with a dummy variable for 1918. These models allow excess pandemic mortality to vary according to baseline population health. The analysis relies on four different measures of city health capital: pre-pandemic infant mortality, baseline typhoid mortality (a marker for water quality), population density in 1910, and pre-pandemic city spending for health and sanitation.²⁰

Columns (2) to (5) in Table 3 report the results for infant and all-age mortality.

¹⁹See Mamelund (2006) for a discussion.

²⁰Average pre-pandemic infant mortality and city spending are calculated for the years 1915 to 1917. Typhoid mortality is measured as the average annual deaths per 100,000 population for the period 1900 to 1905, compiled from Whipple (1908). Information on city spending on health and sanitation is from Miller (2008).

For reference, column (1) reports the baseline estimates. The effects of coal-fired capacity on mortality are robust to the various health capital controls. For both infant and all-age mortality, the estimates are quite stable across specifications and are all highly significant.

Table 3 shows some evidence that lower city-level health capital, as measured by pre-pandemic infant mortality and baseline typhoid mortality, contributed to all-age pandemic severity.²¹ To further explore the relationship between pollution, health capital, and the pandemic, we estimate the interaction effect between coal-fired capacity and pandemic severity across each tercile of pre-pandemic infant mortality.²² This analysis allows us to directly examine the pollution-pandemic relationship across different levels of baseline health capital. By estimating the impact of coal-fired capacity on pandemic-related mortality across groups of cities with similar levels of baseline health, we can further address the issue of selection of less healthy individuals into heavily polluted cities.

Table 4 reports these results. For both infant and all-age mortality, the estimated effects of coal capacity on pandemic-severity are significant in all but the lowest tercile. The estimated pollution-pandemic interaction effects increase monotonically with baseline mortality.²³ For example, the estimated interaction effects for infants are 0.0209, 0.0297, and 0.0414 across the lower, middle, and upper terciles of baseline mortality. These results suggest that exposure to air pollution interacts with baseline health capital more broadly to make individuals more susceptible to the consequences

²¹The magnitude of these point estimates are not directly comparable to the effects for coal-fired capacity. Rescaling the coefficients by the standard deviation in each variable, we find that the two measures of health capital account for a slightly smaller fraction of the cross-city variation in excess pandemic mortality than differences in coal-fired capacity.

²²Pre-pandemic infant mortality is defined as average city-level infant mortality between 1915 and 1917.

²³Across all four models, F-tests for equality of the coefficients are all rejected at 5 percent significance levels.

of a negative health shock. These findings are consistent with a generalized version of the conceptual framework presented in the Appendix, in which air pollution interacts with an individual’s baseline health endowment to affect pandemic severity. Importantly, the significant effects of pollution on pandemic mortality across cities with comparable baseline levels of health provides strong evidence that the main findings are not driven by the selection of unhealthy individuals into heavily polluted cities.

5.3 Pollution, Timing of Onset, and Pandemic Severity

In this section, we examine whether a spurious correlation between the timing of pandemic onset and local pollution levels can account for the estimated effects reported in Table 2. We interact the 1918 fixed effect with four different measures of the timing of pandemic onset to allow for differences in pandemic mortality based on when the outbreak occurred.

First, we estimate a generalized version of equation (1) that controls for state-year fixed effects. These models rely solely on within-state differences in pandemic severity, which greatly limits the importance of timing, since the onset occurred within a very short time frame within states. Second, we rely on a map by Sydenstricker (1918) to identify the week of approximate onset for each city in the sample. We interact this variable with the 1918 year fixed effect. Third, we interact city distance to the nearest World War I military base with 1918 to allow for differences in mortality based on the spread of the disease driven by wartime troop movements.²⁴ Finally, we interact the pandemic year with the number of miles of railway line per county in 1911 (from Donaldson and Hornbeck, forthcoming), as an additional control for

²⁴The movement of military personnel is believed to have been an important determinant of pandemic timing. Crosby (1989), Kolata (2001), and Byerly (2010) provide detailed accounts of the pandemic in the military, and the role of the Navy and Army in its spread. We digitized information on the location of all major army training camps in 1918 (U.S. War Department, 1919, p.1519) to derive the measures of city distance to bases.

disease transmission.

Columns (2) to (5) in Table 5 report the results from these regressions for infant and all-age mortality. The estimated pollution-pandemic relationship is robust to controls for the timing of pandemic onset. The point estimates are all statistically significant, and all but one are very similar in magnitude to the original estimates. We also find some evidence of relationship between the timing of pandemic onset and mortality. Cities that were hit earlier experienced higher pandemic-related mortality rates. The findings are consistent with previous claims that the virus weakened as it spread across the county (Crosby, 1989).

Table 6 reports the effects of coal-fired capacity on pandemic severity by week of pandemic onset. We estimate a generalized version of equation (1), in which the coal interaction effect is allowed to vary according the timing of initial pandemic onset. Cities are classified into four distinct bins, based on the week of initial onset: Before September 14th, September 14-20, September 21-27, and after September 27. Comparing cities that experienced contemporaneous onset, we again find that coal-fired capacity had substantial effects on pandemic-related mortality. For both infant and all-age mortality, the interaction effects were particularly large in cities hit earlier by the pandemic, and the point estimates are largest among cities that experienced onset prior to September 21.

There are three possible explanations for these results. First, the decreasing effects could be due to local preventative measures that, over time, mitigated the spread of virus. Second, the results could reflect the fact that pollution was particular harmful in communities exposed to the most virulent H1N1 strain. Third, the results could be due to differences in the characteristics of cities according to the timing of pandemic onset. In particular, cities that experienced earlier outbreaks also had

substantially higher levels coal-fired capacity.²⁵ The findings provide further evidence that differences in the timing of onset cannot explain the impact of coal-fired capacity on pandemic mortality.

5.4 Additional Specification Checks

Table 7 examines the robustness of the main findings to several other specifications and controls. For reference, column (1) reports the baseline estimates from column (3) of Table 2. In column (2), we allow for differential trends in pollution across cities, replacing the state-specific trend with a city-specific trend. The results are robust to this alternative specification. In column (3) we restrict the sample to the balanced panel of 129 cities reporting mortality in each year between 1915 and 1925. The point estimates are similar to the baseline results, indicating that non-random sample selection is not driving the original findings.

Columns (4) to (6) examine the sensitivity of the estimates to alternative measures of local pollution exposure. Column (4) reports the results based on coal-fired capacity in levels (100s of megawatts). The results are again sizeable, although the coefficient for all-age mortality is imprecisely estimated. Columns (5) and (6) measure coal-fired capacity within 50 and 100 miles of each city-centroid. The point estimates are again large and statistically significant, confirming that the results are not sensitive to the particular distance cutoff.²⁶ The effect of a one standard deviation change in coal-fired capacity is similar across the three distances. In unreported regression models that include both coal-fired capacity within 30 miles and 100 miles, only the former has a significant impact on pandemic-related mortality. These findings support our

²⁵In unreported regressions, we find that the impact of coal-capacity on pandemic-related mortality increases with capacity, which may explain the heterogeneous pattern reported in Table 6.

²⁶Regressions weighted by city population yield qualitatively similar results. Unreported propensity-score matching estimates also confirm that cities with above-median coal-fired capacity experienced significantly higher pandemic-related mortality rates.

choice of distance, and are consistent with evidence suggesting that the historical dispersion of power plant emissions was highly localized.

6 Quantifying the Impact of Coal-Fired Capacity on Pandemic Mortality

To quantify the role of air pollution in exacerbating the pandemic, we assess the severity of the influenza pandemic in two alternative scenarios. We first consider a setting in which all above-median cities reduced coal-fired capacity to the median level. In this scenario, individuals in above-median cities would experience a mean reduction in coal-fired capacity of 1.77 log points.²⁷ Applying the point estimates for coal-fired capacity, we calculate the number of infant deaths and all-age deaths that would have been averted in this scenario.

Table 8 reports these mortality reductions. Columns (1) and (3) report the number of deaths averted under the assumption that the effect of coal capacity on pandemic mortality occurred solely in 1918. We calculate that 3,283 infant deaths and 15,617 all-age deaths would have been averted had coal capacity been reduced to the median-city level, a decrease of roughly 36 infant deaths and 170 all-age deaths per city.²⁸ In columns (2) and (4) of Table 8, we allow coal capacity to affect excess mortality in both 1918 and 1919, based on the linear combination of the interaction effects

²⁷To derive the change in individual-level exposure to coal-fired capacity, cities are weighted by city population (births) in 1921. Intuitively, this scenario is equivalent to moving all individuals in above-median coal capacity to a median capacity city. As a result, we want to rely on the change in individual-level exposure rather than the change in city-level exposure when calculating the number of lives saved.

²⁸To derive these estimates, we first calculate the change in death probability associated with the pollution reduction and then multiply by the total population in high coal cities: $\Delta\text{Probability of a death} = \hat{\beta}_{Coal} \cdot \Delta\text{Log}(\text{CoalCap30mile}) \cdot \overline{\text{MortRate}}_{1918}/10,000 = 0.0216 \times 1.77 \times 191.4/10,000 = 0.000734$. Given a total population 21.3 million in above-median cities, the number of deaths averted is calculated to be 15,617.

reported in Table A.3. In this scenario, we calculate that a reduction in coal capacity to the median level would have averted 5,232 infant deaths and 22,314 all-age deaths, a decrease of 57 infant deaths and 243 all-age deaths per city.

The results suggest that air pollution played an important role in exacerbating the pandemic. By reducing coal capacity to the median level, above median cities would have experienced a 25-40 percent decrease in pandemic-related infant mortality and a 10-14 percent decrease in all-age pandemic-related mortality.²⁹ The large effects found for infants are consistent with early-life health outcomes being particularly sensitive to environmental air quality (Currie et al, 2014). These results also suggest a link between contemporaneous pollution exposure and pandemic severity, given that infant health is more heavily influenced by contemporaneous pollution.³⁰

The previous scenario provided insight into the impact of air pollution on pandemic mortality. In practice, however, it would have been difficult to achieve these reductions in coal-fired capacity. Hydroelectricity was an alternative to coal-fired power, although the feasibility of this substitute depended crucially on local topography. We next consider a setting in which hydroelectric capacity added between 1920 and 1930 was instead been installed prior to the pandemic. This scenario has the advantage that it simply alters the *timing* of capacity installations, and does not impose implausible assumptions on the local availability of hydroelectric potential. We digitized information on the location of hydroelectric power plants built between 1920 and 1930 (Federal Power Commission, 1962). For each city in the sample, we calculate the change in hydroelectric capacity within 30 miles between 1920 and 1930. We then

²⁹Pandemic-related deaths are derived by combining the estimates of excess mortality in 1918 with total population (number of infants) in above-median cities.

³⁰The fact that pre-1915 air quality is unobserved will also tend to downward bias the estimates for all-age mortality. Currie et al (2014) show that exposure to pollution in early childhood is particularly important for later health outcomes. The large rural outmigration that occurred during the late 19th and early 20th century implies that many city residents were not exposed to urban air pollution during childhood.

calculate the required coal-fired capacity necessary to maintain energy production at its 1918 level.

This scenario is associated with an average city-level decrease in coal-fired capacity of 0.275 log points. Combining the health interaction effects in 1918 and 1919, we calculate that 998 infant deaths and 4,630 all-age deaths could have been averted had coal capacity been replaced with readily available hydropower. These estimates correspond to a 7 percent decrease in pandemic-related infant mortality and a 2 percent decrease in all-age pandemic-related mortality.

7 Conclusion

This paper provides new evidence on the extent to which air pollution affected infant and all-age mortality during 1918-1919 influenza pandemic. The effects of air pollution were sizeable. Cities with above median levels of coal-fired capacity collectively experienced thousands of excess infant deaths and tens of thousands of excess all-age deaths during the pandemic. Pre-pandemic health conditions also contributed to pandemic severity, and the marginal impact of coal was larger in places with lower baseline health capital.

The 1918-1919 influenza pandemic was an exceptional episode, with death rates 5 to 20 times higher than subsequent pandemics. Nevertheless, even a moderately severe modern pandemic could lead to 2 million excess deaths (Fan, Jamison, and Summers, 2016), and a pandemic virus with similar pathogenicity to the 1918-1919 virus would likely kill more than 100 million people worldwide (Taubenberger and Morens, 2006). There is a serious risk that air pollution could magnify these health costs, given that a far greater share of the global population live in heavily polluted cities than did in the early 20th century. Our findings highlight the need for research

on the impact of air pollution on later pandemics, including the 1957-1958, 1968-1969, and 2008-2009 pandemics, when healthcare was more readily available. Even with modern antiviral and antibacterial drugs, a global pandemic would quickly overwhelm existing medical infrastructure. Thus, preventative approaches including pollution abatement, and strategic allocation of vaccination efforts are likely to be critical for mitigating mortality.

References

Acuna-Soto, Rodolfo, Cecilia Viboud, and Gerardo Chowell. 2011. "Influenza and pneumonia mortality in 66 large cities in the United States in years surrounding the 1918 pandemic." *PloS one*, 6(8): e23467.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-term Effects of In Utero Infuenza Exposure in the Post-1940 U.S. Population", *Journal of Political Economy*, 114(4): 672-712.

Almond, Douglas, Yuyu Chen, Michael Greenstone, and Hongbin Li. 2009. "Unintended Consequences of China's Huai River Policy," *American Economic Review: Papers and Proceedings*, 99: 184-190.

Almond, Douglas, and Janet Currie. 2011. "Human Capital Development Before Age Five." In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Vol. 4B: 1315-1486. Amsterdam: Elsevier.

Arceo-Gomez, Eva, Rema Hanna, and Paulina Oliva. 2012. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." NBER Working Paper #18349.

Bootsma, Martin, and Neil Ferguson. 2007. "The effect of public health measures on the 1918 influenza pandemic in U.S. cities." *PNAS*, 104(18): 7588-7593.

Brainerd, Elizabeth, and Mark V. Siegler. 2003. "The Economic Effects of the 1918 Influenza Epidemic." Discussion Paper no. 3791, Centre Econ. Policy Res., Paris.

Brook, Robert, et al. 2010. "Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association," *Circulation*, 121: 2331-2378.

Byerly, Carol. 2010. "The U.S. Military and the Influenza Pandemic of 1918-1919," *Public Health Reports*, 125(3): 82-91.

Cauchemez, Simon, Alain-Jacques Valleron, Pierre-Yves Boelle, Antoine Flahaut, and Neil Ferguson. 2010. “Estimating the impact of school closure on influenza transmission from Sentinel data,” *Nature*, 452(7188): 750-755.

Chay, Kenneth Y. and Michael Greenstone. 2003a. “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession.” *Quarterly Journal of Economics*, 118: 1121-1167.

Chay, Kenneth Y. and Michael Greenstone. 2003b. “Air Quality, Infant Mortality, and the Clean Air Act of 1970.” MIT Department of Economics Working Paper No. 04-08.

Chowell, Gerardo, et al. 2008. “The 1918-1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact.” *Proceedings of the Royal Society B: Biological Sciences*, 275.1634: 501-509.

Cohen, Aaron et al. 2004. “Chapter 17: Urban Air Pollution” in *Comparative Quantification of Health Risks, Volume 2*. Geneva: World Health Organization.

Collins, Selwyn D. 1930. “Age and Sex Incidence of Influenza and Pneumonia Morbidity and Mortality in the Epidemic of 1928-29 with Comparative Data for the Epidemic of 1918-19: Based on Surveys of Families in Certain Localities in the United States following the Epidemics,” *Public Health Reports*, 46(33): 1909-1937.

Crosby, Alfred D. 1989. *America’s Forgotten Pandemic: The Influenza of 1918*. New York: Cambridge Univ. Press.

Currie, Janet, Joshua Graff-Zivin, Jamie Mullen, and Matthew Neidell. 2014. “What Do We Know About Short and Long Term Effects of Early Life Exposure to Pollution?” *Annual Review of Resource Economics*, 6: 217-247.

Currie, Janet and Matthew Neidell. 2005. “Air Pollution and Infant Health: What Can We Learn From California’s Recent Experience?” *Quarterly Journal of Economics*, 120: 1003-1030.

Currie, Janet, and Reed Walker. 2011. “Traffic Congestion and Infant Health: Evidence from E-Z Pass,” *American Economic Journal: Applied Economics*, 3(1): 65-90.

Delfino, Ralph J., Constantinos Sioutas, and Shaista Malik. 2005. “Potential Role of Ultrafine Particles in Associations between Airborne Particle Mass and Cardiovascular Health,” *Environmental Health Perspectives*, 113: 934-946.

Donaldson, Dave, and Richard Hornbeck. Forthcoming. “Railroads and American Economic Growth: A “Market Access” Approach,” *Quarterly Journal of Economics*.

Eisenbud, Merrill. 1978. Levels of Exposure to Sulfur Oxides and Particulates in New York City and their Sources. *Bulletin of the New York Academy of Medicine* 1978, 54:991-1011.

Fan, Victoria, Dean T. Jamison, and Lawrence H. Summers. 2016. "The Inclusive Cost of Pandemic Influenza Risk," NBER Working Paper #22137.

Flagg, Samuel B. 1912. *City Smoke Ordinances and Smoke Abatement*. Washington D.C.: Government Printing Office.

Goklany, Indur M. *Clearing the air: the real story of the war on air pollution*. Cato Institute, 1999.

Goodkind, Andrew L., Jay S. Coggins, and Julian D. Marshall. 2014. "A Spatial Model of Air Pollution: The Impact of the Concentration-Response Function," *Journal of the Association of Environmental and Resource Economists*, 1(4): 451-479.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2): 223-255.

Hahon, Nicholas, et al. 1985. "Influenza virus infection in mice after exposure to coal dust and diesel engine emissions," *Environmental Research*, 37(1): 44-60.

Haines, Michael R., and Inter-university Consortium for Political and Social Research (ICPSR). 2010. *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, icpsr.org.

Hales, Jeremy M. 1976. "Tall Stacks and the Atmospheric Environment," *EPA Publication No. EPA-450/3-76-007*.

Harrod, Kevin, et al. 2003. "Increased Susceptibility to RSV Infection by Exposure to Inhaled Diesel Engine Emissions," *American Journal of Respiratory Cell and Molecular Biology*, 28(4): 451-463.

Hoek, Gerard, et al. 2013. "Long-term air pollution exposure and cardio-respiratory mortality: a review," *Environmental Health*, 12(43): 1-15.

Huntington, Ellsworth. 1923. "Causes of Geographical Variation in the Influenza Epidemic in the Cities of the United States," *Bull. Nat. Res. Council*, 6: 1-36.

Ireland, Merritt W (Ed.). 1928. "Medical Department of the United States Army in the World War," *Communicable Diseases*, 9: Washington, D.C.: U.S. Army.

Ives, James E., et al. 1936. *Atmospheric Pollution of American Cities for the Years 1931 to 1933 with Special Reference to the Solid Constituents of the Pollution*. U.S. Treasury Department, Public Health Bulletin No 224. Washington: Government Printing Office.

Jakab, George J. 1993. "The Toxicological Interactions Resulting from Inhalation of Carbon Black and Acrolein on Pulmonary Antibacterial and Antiviral Defenses," *Toxicology and Applied Pharmacology*, 121: 167-175.

Jaspers, Ilona, et al. 2005. "Diesel Exhaust Enhances Influenza Virus Infections in Respiratory Epithelial Cells," *Toxicology Sciences*, 85(2): 990-1002.

Johnson, Niall. and Juergen Mueller. 2002. *Bulletin of Historical Medicine*, 76: 105-115.

Kolata, Gina. 1999. *Flu: The Story of the Great Influenza Pandemic of 1918 and the Search for the Virus That Caused It*. New York: Touchstone.

Lee, Greg, et al. 2014. "Exposure to combustion generated environmentally persistent free radicals enhances severity of influenza virus infection," *Particle and fibre toxicity*, 11(1): 57.

Lelieveld, J., J. S. Evans, M. Fnais, D. Giannadaki, and A. Pozzer. 2015. The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale, *Nature*, 525: 367-371.

Levy, Jonathan I., et al. 2002. "Using CALPUFF to evaluate the impacts of power plant emissions in Illinois: model sensitivity and implications," *Atmospheric Environment*, 36: 1063-1075.

Lin, Ming-Jen, and Elaine M. Liu. 2014. "Does in *utero* exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment," *Journal of Health Economics*, 37: 152-163.

Mamelund, Sverre-Erik. 2006. "A Socially Neutral Disease? Individual Social Class, Household Wealth and Mortality from Spanish Influenza in Two Socially Contrasting Parishes in Kristiania 1918-19," *Social Science and Medicine* 62: 923-40.

Meltzer, Martin I., Nancy J. Cox., and Keiji Fukuda. 2005. "The Economic Impact of Pandemic Influenza in the United States: Priorities for Intervention," In Knobler, Stacey L., Alison Mack, Adel Mahmoud, and Stanley M. Lemon (eds.) *The Threat of Pandemic Influenza: Are We Ready? Workshop Summary*, National Academies Press.

Miller, Grant. "Women's Suffrage, Political Responsiveness, and Child Survival in American History," *Quarterly Journal of Economics*, 123(3): 1287-1327.

- Neelsen, Sven, and Thomas Stratmann. 2012. "Long-Run Effects of Fetal Influenza Exposure: Evidence from Switzerland," *Social Science and Medicine*, 74(1): 58-66.
- Nelson, Richard E. 2010. "Testing the fetal origins hypothesis in a developing country: Evidence from the 1918 influenza pandemic," *Health Economics*, 19(10): 1181-1192.
- Pearl, Raymond. 1921. "Influenza studies: further data on the correlation of explosiveness of outbreak of the 1918 epidemic." *Public Health Report*, 36: 273-298.
- Pope, C. Arden III, et al. 2004. "Cardiovascular mortality and longterm exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease," *Circulation*, 109: 71-77.
- Pope, C. Arden III, Maureen Cropper, Jay Coggins, and Aaron Cohen. 2015. "Health Benefits of Air Pollution Abatement Policy: Role of the Shape of the Concentration-Response Function," *Journal of the Air & Waste Management Association*, 65(5): 516-522.
- Seinfeld, John H., and Spyros N. Pandis. 2012. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons.
- Stern, Arthur C. "History of Air Pollution Legislation in the United States." *Journal of the Air Pollution Control Association* 32.1 (1982): 44-61.
- Sydenstricker, Edgar. 1918. "Preliminary Statistics of the Influenza Epidemic." *Public Health Reports* 33 (December): 2305-21.
- Taubenberger, Jeffery, and David M. Morens. 2006. "1918 Influenza: the Mother of All Pandemics," *Emerging Infectious Diseases*, 12(1): 15-22.
- U.S. Census Bureau. 1919. *Fourteenth Census of the United States, Volume XI Mines and Quarries, General Report and Analytical Tables and Selected Industries*. Washington D.C.: U.S. Census Bureau.
- U.S. Census Bureau. 1976. *Historical Statistics of the United States, Colonial Times to Present*. Washington D.C.: U.S. Census Bureau.
- U.S. Department of Agriculture. 1916. *Electric Power Development in the United States*. Washington D.C.: U.S. Government Printing Office.
- U.S. Geological Survey. 1917. *Mineral Resources of the United States. Part II – Non-metals*. Washington D.C.: U.S. Government Printing Office.
- U.S. National Office of Vital Statistics, *Vital Statistics of the United States*, various year. Washington D.C.: U.S. Government Printing Office.

U.S. Federal Power Commission. 1962. *Principal Electric Power Facilities in the United States (map)*. Washington D.C.: U.S. Federal Power Commission.

U.S. War Department. 1919. *Second report of the Provost Marshall General to the Secretary of War on the operations of the Selective Service System to December 20, 1918*. Washington, D.C.: U.S. Government Printing Office.

Whipple, George, C. 1908. *Typhoid Fever: Its Causation, Transmission, and Prevention*. New York: John Wiley and Sons.

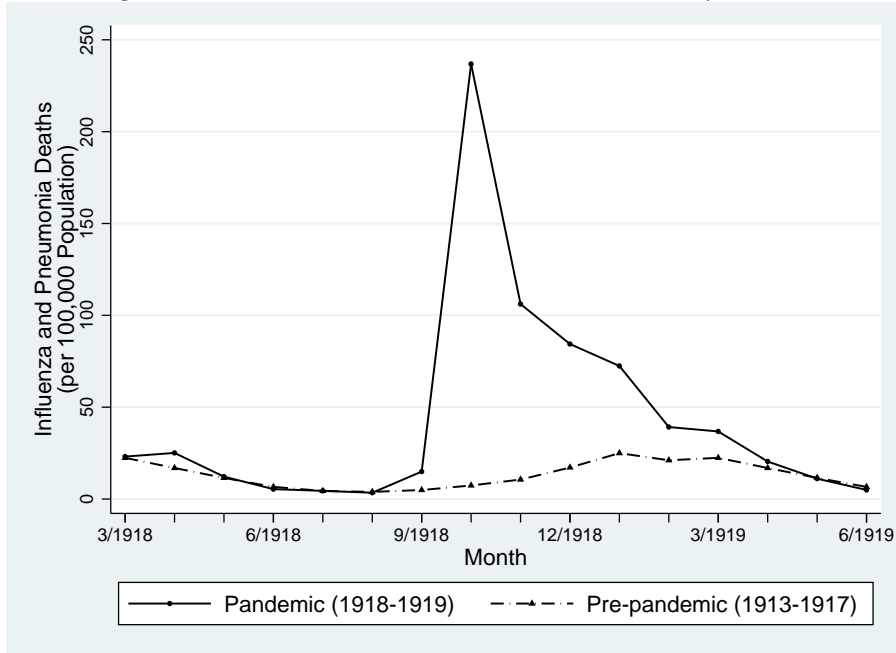
World Bank. 2013. *World Development Report: Pandemic Risk*. <https://openknowledge.worldbank.com/handle/10986/16343>.

World Health Organization. 2014. *World Health Statistics: 2014*. http://www.who.int/gho/publications/world_health_statistics/en/ (accessed September 2015).

Woodruff, Tracey, Lyndsey Darrow, and Jennifer Parker. 2008. "Air Pollution and Postneonatal Infant Mortality in the United States, 1999-2002." *Environmental Health Perspectives* 116: 110-115.

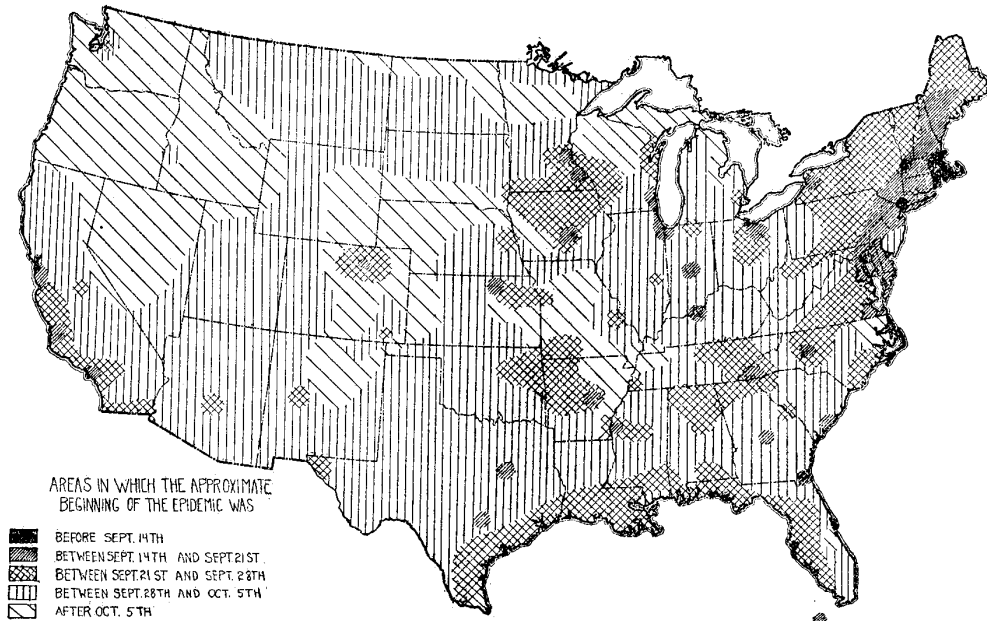
Xu, Zhiwei, et al. 2013. "Air pollution, temperature, and pediatric influenza in Brisbane, Australia," *Environmenta International*, 59: 384-388.

Figure 1: Influenza and Pneumonia Deaths by Month



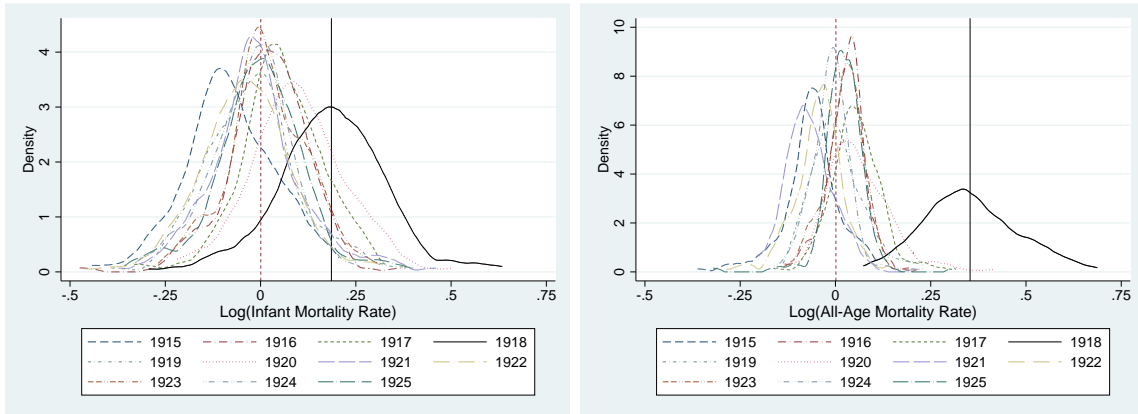
Source: Vital Statistics, 1913-1919.

Figure 2: Timing of Pandemic Onset for 376 Localities



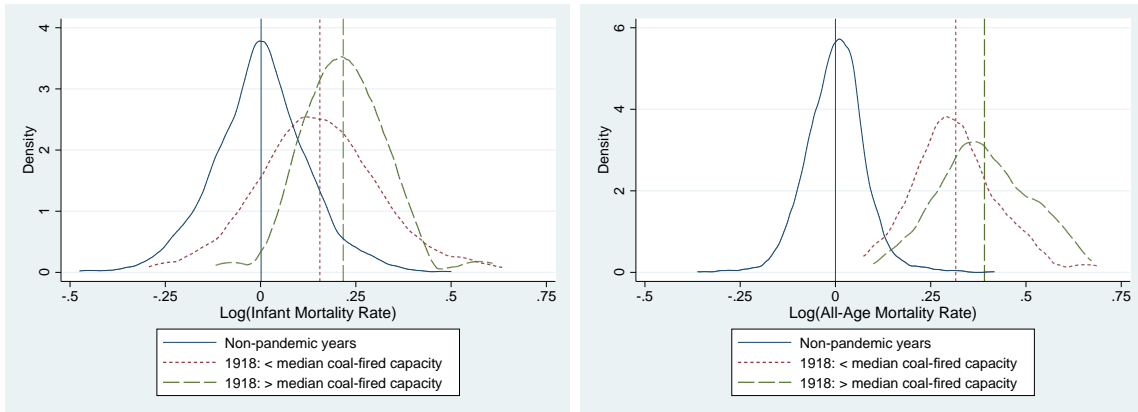
Source: Sydenstricker (1918).

Figure 3: Excess Mortality by Year



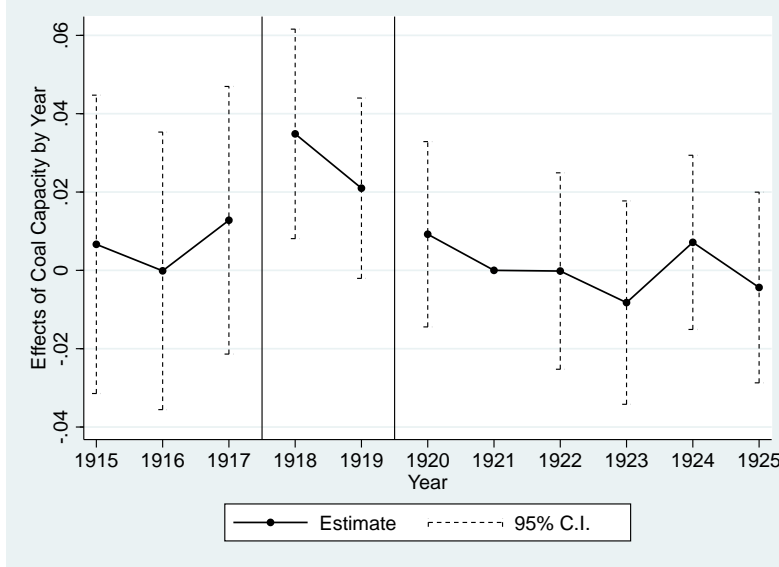
Notes: This figure reports the density of residual infant mortality and residual all-age mortality from regressions that controls for a linear city-specific trend. Mortality is plotted separately for each year in the sample between 1915 and 1925.

Figure 4: Excess Mortality in 1918 for Cities Above and Below Median Coal-Fired Capacity



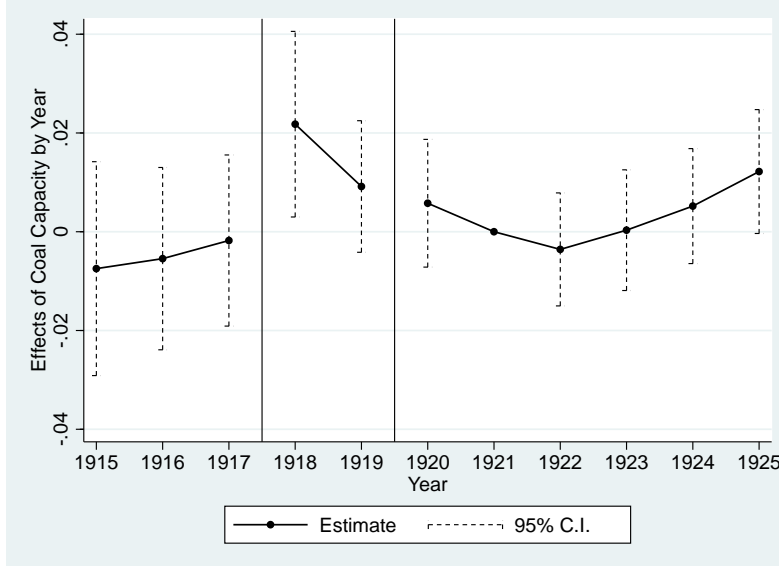
Notes: This figure reports the density of residual mortality from a regression that controls for a linear city-specific trend. The distributions are plotted in non-pandemic years and in pandemic years for cities above and below median coal-fired capacity.

Figure 5: Effect of Coal-Fired Capacity on Infant Mortality, by Year



Notes: This figure reports the vector of coal capacity-year interaction effects based on the regression model estimated in column (4) of Table 2. These models are a generalized version of equation (1), that allows coal-fired capacity to affect mortality in all years. The solid line denotes the estimates, and the dotted lines denote the 95% confidence interval.

Figure 6: Effect of Coal-Fired Capacity on All-Age Mortality, by Year



Notes: This figure reports the vector of coal capacity-year interaction effects based on the regression model estimated in column (4) of Table 2. These models are a generalized version of equation (1), that allows coal-fired capacity to affect mortality in all years. The solid line denotes the estimates, and the dotted lines denote the 95% confidence interval.

Table 1: Summary Statistics

	Mean	Standard deviation
<i>Panel A: Mortality</i>		
Log(infant mortality)	4.420	0.279
Log(all-age mortality)	7.198	0.251
<i>Panel B: Hydroelectric and coal-fired capacity</i>		
Log(coal-fired capacity with 30 miles)	3.837	1.865
Coal-fired capacity within 30 miles	181.4	325.0
Hydro capacity within 30 miles	11.60	29.82
<i>Panel C: City and County characteristics</i>		
City population in 1921	155,801	468,398
City births 1921	3,780	11,005
Population density in 1910	1,571	4,497
Share urban in 1910	0.756	0.167
Share white in 1910	0.950	0.105
Log(manufacturing payroll per population in 1900)	0.959	0.049
Employment share in manufacturing in 1910	0.143	0.059
<i>Panel D: Other determinants of pandemic severity</i>		
Pre-pandemic infant mortality rate	92.8	28.3
Pre-pandemic typhoid mortality rate	35.8	24.4
Pre-pandemic health expenditure (per capita)	1.31	1.00
Timing of pandemic onset		
Pre Sept. 14	0.20	0.40
Sept. 14 - Sept. 21	0.29	0.45
Sept. 21 - Sept. 28	0.22	0.42
Post Sept. 28	0.29	0.45
Distance to nearest WWI base	87.6	84.0
Miles of railroad lines in 1911	121.2	77.1
Number of cities	183	
Observations	1769	

Notes: Column 1 reports the sample means. Column 2 reports standard deviation. Pre-pandemic infant mortality rate is calculated as the average city-level infant mortality rate (per 1,000 live births) for the years 1915-1917. Pre-pandemic typhoid mortality rate is calculate as the average city-level typhoid mortality rate (per 100,000 population) for the years 1900-1905.

Table 2: The Effect of the Pandemic on Mortality, by Coal-Fired and Hydroelectric Capacity

	Estimated effect			
	(1)	(2)	(3)	(4)
Dependent variable: Log(infant mortality)				
I(Year=1918) \times Log(CoalCap30mile)	0.0288*** [0.00584]	0.0286*** [0.00702]	0.0309*** [0.00829]	0.0349** [0.0136]
I(Year=1918) \times Log(HydroCap30mile)	0.00001 [0.00627]	-0.00233 [0.00806]	0.000589 [0.00891]	0.0128 [0.0123]
P-value from test of $\beta_{Coal} = \beta_{Hydro}$	<0.001	0.002	0.003	0.106
Observations	1,769	1,769	1,769	1,769
Number of clusters	183	183	183	183
Sample mean				
Log(excess infant mortality in 1918)	Mean = 0.186, S.D. = 0.143			
Dependent variable: Log(all-age mortality)				
I(Year=1918) \times Log(CoalCap30mile)	0.0218*** [0.00541]	0.0233*** [0.00735]	0.0216*** [0.00782]	0.0219** [0.00954]
I(Year=1918) \times Log(HydroCap30mile)	0.00145 [0.00563]	0.00210 [0.00613]	0.00314 [0.00687]	0.00583 [0.00790]
P-value from test of $\beta_{Coal} = \beta_{Hydro}$	0.007	0.027	0.052	0.128
Observations	1,768	1,768	1,768	1,768
Number of clusters	183	183	183	183
Sample mean				
Log(excess all-age mortality in 1918)	Mean = 0.352, S.D. = 0.123			
City & year FE + Linear state trend	Y	Y	Y	Y
Demographic & geographic covariates \times year		Y	Y	Y
Annual city-level health expenditure			Y	Y
Manufacturing covariates \times year			Y	Y
Coal & hydro capacity \times year				Y

Notes: Each column reports the point estimates from a different regression. The variables Log(CoalCap30mile) and Log(HydroCap30mile) denote logarithm of coal-fired and hydroelectric capacity within 30 miles of the city-centroid. Demographic and geographic covariates include county-level controls for the logarithm of population density, percent urban, and share white in 1910, and longitude and latitude. Manufacturing covariates include the logarithm of manufacturing wages in 1900 and the logarithm of manufacturing employment in 1910. The final column includes the full interaction effects of Log(CoalCap30mile) and Log(HydroCap30mile) with year fixed effects. Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 3: Coal-Fired Capacity, Health Capital, and Pandemic Mortality

	Estimated effect				
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Log(infant mortality)					
I(Year=1918) ×					
Log(CoalCap30mile)	0.0309*** [0.00829]	0.0310*** [0.00823]	0.0352*** [0.0101]	0.0309*** [0.00829]	0.0265*** [0.00711]
<i>Health Capital Covariates</i>					
I(Year=1918) ×					
Log(Baseline infant mortality)		0.0547 [0.0533]			
Log(Baseline typhoid mortality)			0.0295 [0.0209]		
Log(Population density)				0.0169 [0.0188]	
Log(Pre-1918 health expenditure)					0.0146 [0.0135]
Observations	1,769	1,762	1,442	1,769	1,679
Number of clusters	183	179	141	183	156
Dependent variable: Log(all-age Mortality)					
I(Year=1918) ×					
Log(CoalCap30mile)	0.0216*** [0.00782]	0.0218*** [0.00723]	0.0211** [0.00829]	0.0216*** [0.00782]	0.0163** [0.00713]
<i>Health Capital Covariates</i>					
I(Year=1918) ×					
Log(Baseline infant mortality)		0.129*** [0.0482]			
Log(Baseline typhoid mortality)			0.0519*** [0.0189]		
Log(Population density)				0.00942 [0.0132]	
Log(Pre-1918 health expenditure)					-0.00563 [0.0143]
Observations	1,768	1,761	1,441	1,768	1,678
Number of clusters	183	179	141	183	156

Notes: All models include the full set of controls reported in column (3) of Table 2. Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 4: Heterogeneous Effects by Pre-Pandemic Mortality

	Estimated effect			
	(1)	(2)	(3)	(4)
	Dependent variable: Log(infant mortality)		Dependent variable: Log(all-age mortality)	
<i>Heterogeneity by tercile of pre-pandemic infant mortality</i>				
I(Year=1918) × Log(CoalCap30mile) ×				
Lower tercile	0.0156** [0.0076]	0.0209** [0.0091]	0.0016 [0.0073]	0.0067 [0.0083]
Middle tercile	0.0263*** [0.0063]	0.0297*** [0.0077]	0.0189*** [0.0053]	0.0216*** [0.0065]
Upper tercile	0.0354*** [0.0063]	0.0414*** [0.0089]	0.0313*** [0.0064]	0.0351*** [0.0080]
P-value from test: $\beta_{Coal}^L = \beta_{Coal}^M = \beta_{Coal}^H$	0.005	0.011	0.000	0.001
Observations	1,769	1,769	1,768	1,768
Number of clusters	183	183	183	183
City & Year FE + Linear state trend	Y	Y	Y	Y
Full controls		Y		Y

Notes: Each column reports the point estimates from a different regression. Columns (2) and (4) include the full set of controls reported in column (3) of Table 2. The table reports the interaction effects across the three terciles of pre-pandemic infant mortality (between 1915 and 1917). Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 5: Coal-Fired Capacity, Timing of Onset, and Pandemic Mortality

	Estimated effect				
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Log(infant mortality)					
I(Year=1918) ×					
Log(CoalCap30mile)	0.0309*** [0.00829]	0.0352*** [0.00987]	0.0254*** [0.00845]	0.0281*** [0.00808]	0.0300*** [0.00831]
<i>Pandemic Timing Covariates</i>					
State-year fixed effects		Yes			
I(Year=1918) ×					
Week of pandemic onset			-0.0255* [0.0139]		
Log(Distance to WWI base)				-0.0279* [0.0168]	
Log(Railway lines)					0.0134 [0.0155]
Observations	1,769	1,769	1,769	1,769	1,769
Number of clusters	183	183	183	183	183
Dependent variable: Log(all-age Mortality)					
I(Year=1918) ×					
Log(CoalCap30mile)	0.0216*** [0.00782]	0.0245*** [0.00926]	0.0151* [0.00788]	0.0217*** [0.00792]	0.0205*** [0.00771]
<i>Pandemic Timing Covariates</i>					
State-year fixed effects		Yes			
I(Year=1918) ×					
Week of pandemic onset			-0.0303** [0.0149]		
Log(Distance to WWI base)				0.00100 [0.0137]	
Log(Railway lines)					0.0163 [0.0115]
Observations	1,768	1,768	1,768	1,768	1,768
Number of clusters	183	183	183	183	183

Notes: All models include the full set of controls reported in column (3) of Table 2. Week of pandemic onset takes values from 0 to 4, increasing in the time to pandemic. Railway lines denote the miles of railway lines within the county. Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 6: Heterogeneous Effects by Week of Pandemic Onset

	Estimated effect			
	(1)	(2)	(3)	(4)
	Dependent variable: Log(infant mortality)		Dependent variable: Log(all-age mortality)	
<i>Heterogeneity by week of pandemic onset</i>				
I(Year=1918) × Log(CoalCap30mile) ×				
I(Pre Sept. 14)	0.0245*** [0.0064]	0.0303*** [0.0083]	0.0146*** [0.0054]	0.0176** [0.0074]
I(Sept. 14 - Sept. 21)	0.0296*** [0.0065]	0.0332*** [0.0086]	0.0257*** [0.0064]	0.0275*** [0.0078]
I(Sept. 21 - Sept. 28)	0.0159* [0.0089]	0.0198** [0.0096]	0.0124 [0.0107]	0.0153 [0.0113]
I(Post Sept. 28)	0.0092 [0.0094]	0.0144 [0.0098]	-0.00213 [0.0090]	0.0012 [0.0101]
P-value from test:				
Equality of effects	0.034	0.083	0.004	0.025
Observations	1,769	1,769	1,768	1,768
Number of clusters	183	183	183	183
City & Year FE + Linear state trend	Y	Y	Y	Y
Full controls		Y		Y

Notes: Each column reports the point estimates from a different regression. Columns (2) and (4) include the full set of controls reported in column (3) of Table 2. The table reports the interaction effects by week of pandemic onset based on Sydenstricker (1918). Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 7: Robustness Checks

	Baseline (1)	Control for city-level trend (2)	Estimated effect			
			Balanced panel (3)	Capacity (100s of MW) (4)	Alternative measures of capacity Within 50 miles (5) Within 100 miles (6)	
Dependent variable: Log(infant mortality)						
I(Year=1918) ×						
Log(CoalCap30mile)	0.0309*** [0.0083]	0.0273*** [0.0079]	0.0228*** [0.0085]	0.00668* [0.0036]	0.0336*** [0.0095]	0.0324** [0.0136]
Observations	1,769	1,769	1,375	1,769	1,769	1,769
Number of clusters	183	183	125	183	183	183
Dependent variable: Log(all-age mortality)						
I(Year=1918) ×						
Log(CoalCap30mile)	0.0216*** [0.0078]	0.0214*** [0.0068]	0.0176** [0.0085]	0.00545 [0.0043]	0.0307*** [0.0088]	0.0268** [0.0122]
Observations	1,768	1,768	1,364	1,768	1,768	1,768
Number of clusters	183	183	124	183	183	183

Notes: Each cell reports the point estimates from a different regression. All models include the full set of controls reported in column (3) of Table 2. Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

Table 8: Pandemic-Related Deaths Averted by Reducing Coal-Fired Capacity in Cities

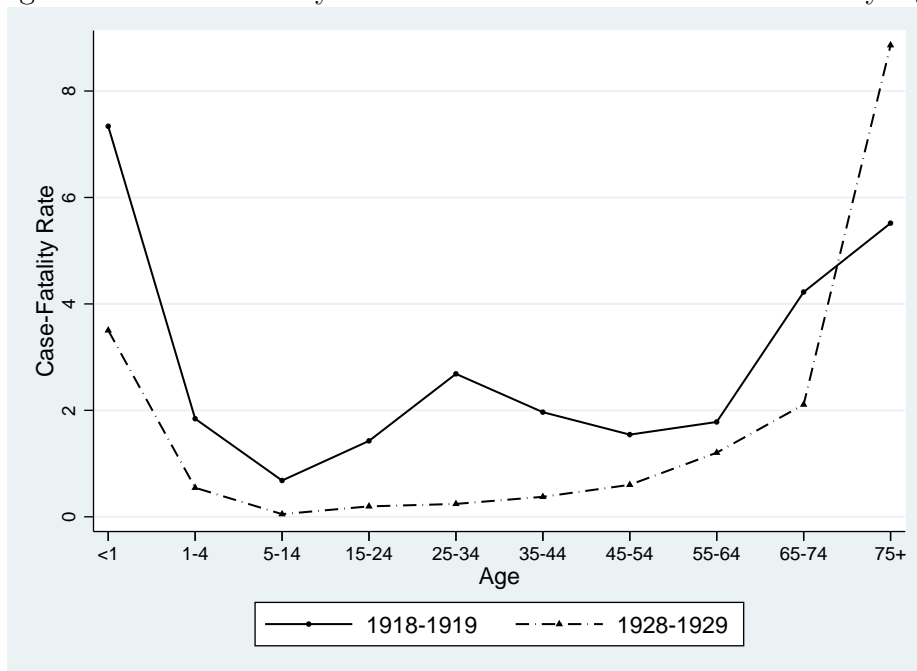
	Infant deaths averted		All-age deaths averted	
	Health impact only in 1918 (1)	Health impact in 1918 & 1919 (2)	Health impact only in 1918 (3)	Health impact in 1918 & 1919 (4)
Scenario 1: <i>Reduce coal-fired capacity in above-median cities to median level (51.7mw)</i>				
# deaths averted	3,283 [880]	5,232 [1,318]	15,617 [5,658]	22,314 [7,668]
Excess mortality in 1918 (above-median cities)		12,973		159,188
Scenario 2: <i>Replace coal-fired capacity with available hydroelectric capacity</i>				
# deaths averted	626 [168]	998 [251]	3,240 [1,172]	4,630 [1,592]
Excess mortality in 1918 (all sample cities)		14,863		192,840

Notes: Each column reports the number of deaths averted under each scenario, based on the regression estimates reported in columns (2) and (4) of Table 3. The number of deaths averted in columns (1) and (3) are derived solely from the 1918 interaction effect. The number of deaths averted in columns (2) and (4) are derived based on the linear combination of the effects in 1918 and 1919. Square brackets report the standard errors for the number of deaths averted. Excess mortality in 1918 is calculated as residual mortality in 1918 after controlling for a city-specific linear trend in mortality.

A Appendix

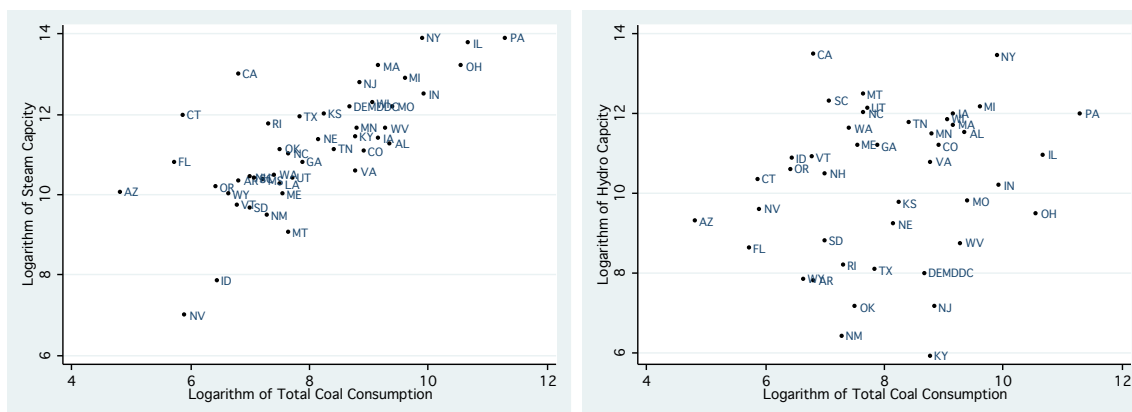
A.1 Figures and Tables

Figure A.1: Case-fatality Rate from Influenza and Pneumonia by Age



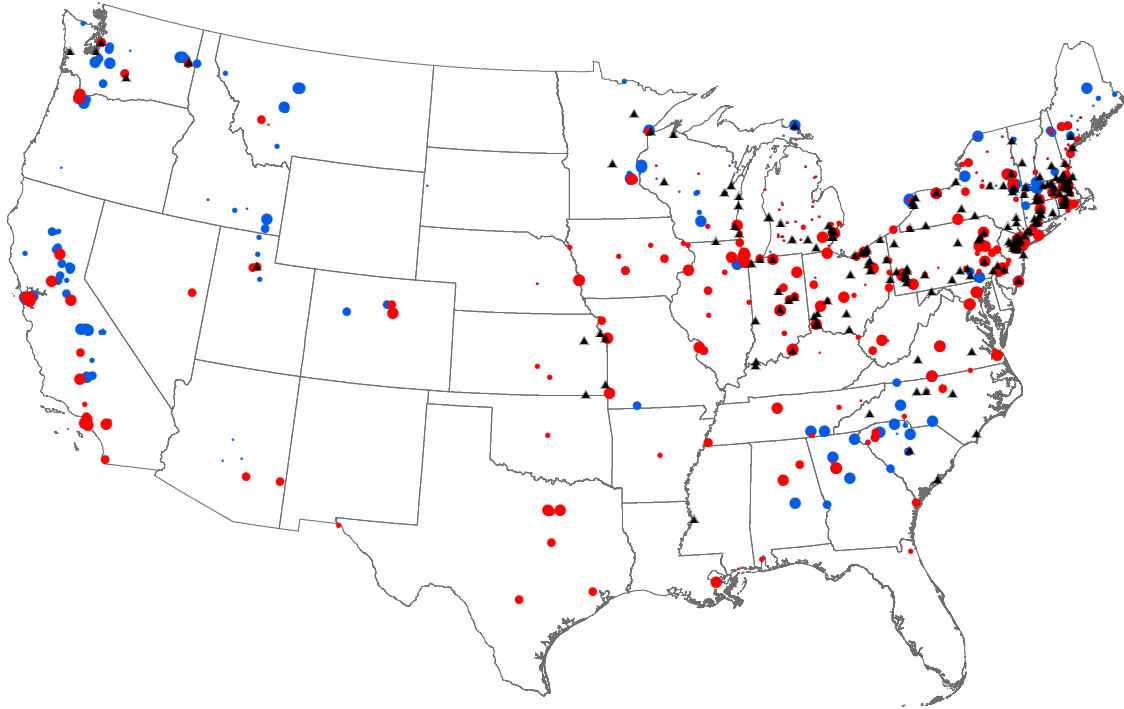
Notes: Based on surveys conducted by the United States Public Health Service in 12 localities in 1918-1919 and 14 localities in 1928-1929. Source: Collins (1931).

Figure A.2: State-level Coal Consumption, Coal-Fired Capacity, and Hydroelectric Capacity



Notes: These figures report the relationship between the logarithm of total state-level coal consumption, coal-fired capacity and hydroelectric capacity in 1917. Source: Mineral Resources of the United States (1917, p.1254).

Figure A.3: Cities in Sample and the Location of Coal-Fired and Hydroelectric Power Plants



Notes: This figure presents the location of 183 cities in the sample, the location of coal-fired and hydroelectric power plants in 1915. Red and blue circles denote steam and hydroelectric power plants by quartile of capacity (<7mw, 7-11mw, 11-18mw, >18mw).

Table A.1: Cities with Smoke Problems and Municipal Smoke Abatement Legislation

Year	Cities with Smoke Problems
1912	<i>Large Cities with Smoke Problems</i> Baltimore, Boston, Buffalo, Chicago, Cincinnati, Cleveland, Denver, Detroit, Indianapolis, Jersey City, Kansas City, Louisville, Milwaukee, Minneapolis, Newark, New York, Philadelphia, Pittsburgh, Providence, Rochester, St. Louis, St. Paul, Washington <i>Large Cities without Smoke Problems</i> Los Angeles, New Orleans, Portland, San Francisco, Seattle
Decade	Cities Passing Smoke Legislation
1880-1890	Chicago, Cincinnati
1890-1900	Cleveland, Pittsburgh, St. Paul
1900-1910	Akron, Baltimore, Boston, Buffalo, Dayton, Detroit, Indianapolis, Los Angeles, Milwaukee, Minneapolis, New York, Newark, Philadelphia, Rochester, St. Louis, Springfield (MA), Syracuse, Washington
1910-1920	Albany County (NY), Atlanta, Birmingham, Columbus, Denver, Des Moines, Duluth, Flint, Hartford, Jersey City, Kansas City, Louisville, Lowell, Nashville, Portland (OR), Providence, Richmond, Toledo

Source: Top: Flag (1912); Bottom: Stern (1982, Table III, p.45).

Table A.2: TSP Concentration in Various Years

Year	Location	TSP
1912-1913	Chicago	760
1931-1933	Baltimore, Boston, Chicago, Pittsburgh, St. Louis	630
1931-1933	Buffalo, Cleveland, New Orleans, New York, Philadelphia	520
1931-1933	Detroit, Los Angeles, San Francisco, Washington	350
1990	US National Average	60
1980-1993	58 Chinese Cities	538
1999	Worldwide	18% urban population > 240

Source: Eisenbud (1978); Ives et al (1936); Chay and Greenstone (2003a); Almond et al (2009); Cohen et al (2004).

Table A.3: The Effect of Coal-Fired Capacity on Mortality, by Year

	Estimated effect			
	(1)	(2)	(3)	(4)
	Dependent variable: Log(infant mortality)		Dependent variable: Log(all-age mortality)	
Log(CoalCap30mile) ×				
I(Year=1915)		0.0067 [0.0193]		-0.0074 [0.0110]
I(Year=1916)		-0.0001 [0.0180]		-0.0054 [0.0094]
I(Year=1917)	0.0091 [0.010]	0.0128 [0.0173]	-0.0021 [0.0064]	-0.0017 [0.0088]
I(Year=1918)	0.0335*** [0.0085]	0.0349** [0.0136]	0.0221*** [0.0083]	0.0219** [0.0095]
I(Year=1919)	0.0179** [0.0087]	0.0210* [0.0117]	0.0083 [0.0051]	0.0092 [0.0068]
I(Year=1920)		0.0091 [0.0118]		0.0057 [0.0066]
I(Year=1921) – Omitted				
I(Year=1922)		-0.0002 [0.0127]		-0.0036 [0.0058]
I(Year=1923)		-0.0082 [0.0132]		0.0003 [0.0062]
I(Year=1924)		0.0072 [0.0113]		0.0052 [0.0059]
I(Year=1925)		-0.0044 [0.0123]		0.0121* [0.0064]
Observations	1,769	1,769	1,768	1,768
Number of clusters	183	183	183	183
Full controls	Y	Y	Y	Y
Coal and hydro capacity × year FE		Y		Y

Notes: Each column reports the point estimates from a different regression. The variable Log(CoalCap30mile) denotes the logarithm of coal-fired capacity within 30 miles of the city-centroid. All models include the full set of controls reported in column (3) of Table 2. Columns (2) and (4) include the full interaction effects of Log(CoalCap30mile) and Log(HydroCap30mile) with year fixed effects. Standard errors are clustered at the city-level. ***, **, * denote significance at the 1%, 5%, and 10%, respectively.

A.2 A Conceptual Framework of the Link between Air Pollution, Health Capital, and Pandemic Severity

To understand the impact of air pollution on mortality during the 1918 Influenza Pandemic, we develop a simple model of health capital that builds on Grossman (1972), Almond and Currie (2011) and Currie et al. (2014). We examine the health capital production function for the average individual in a city. Assume that baseline pre-pandemic health capital, H_B depends on baseline pollution exposure, P_B :

$$H_B = f_B(P_B). \quad (\text{A.1})$$

The impact of pollution on average baseline health capital could be positive or negative. On the one hand, the harmful effects of air pollution will tend to lower health capital throughout the population, implying a decrease in average health capital: $\frac{\partial H_B}{\partial P_B} < 0$. Conversely, if pollution led to increased mortality among individuals with “marginal” health, then selective attrition could lead to an increase in the average population health capital, $\frac{\partial H_B}{\partial P_B} > 0$.

Let H denote the health capital during the 1918 pandemic. We assume that H depends upon the severity of the influenza health shock, I , pre-pandemic health capital, H_B , and exposure to pollution in 1918, P , according to the following function:

$$H = f(I, H_B, P). \quad (\text{A.2})$$

Equations (A.1) and (A.2) can be used to derive the interaction between air pollution exposure and pandemic severity. For simplicity we assume city-level pollution is constant so that $P_B = P$, and take the cross-partial derivative with respect to I and P :

$$\frac{\partial^2 H}{\partial I \cdot \partial P} = f_{I, H_B} \cdot \frac{\partial H_B}{\partial P} + f_{I, P} \quad (\text{A.3})$$

Equation (A.3) shows that city-level pollution affects pandemic severity through two distinct channels. The first term, $f_{I, H_B} \cdot \frac{\partial H_B}{\partial P}$ captures the longer-term impact of pollution exposure on pandemic severity. This effect depends on both the impact of pre-pandemic pollution exposure on average health capital, $\frac{\partial H_B}{\partial P}$, and the relationship between health capital and pandemic severity, f_{I, H_B} . If pollution lowered baseline health and less healthy individuals were more negatively affected by the pandemic, then we should expect the pandemic to have been more severe in heavily polluted areas. The second term captures the relationship between contemporaneous pollution and the pandemic for population health. This effect captures the extent to which short-term pollution exposure increased susceptibility to the influenza virus. The relative magnitude of these two interaction effects is relevant for evaluating the effectiveness of short-term policy interventions, such as temporary restrictions on pollution for mitigating pandemic severity.