

## When Externalities Collide: Influenza and Pollution<sup>†</sup>

By JOSHUA GRAFF ZIVIN, MATTHEW NEIDELL,  
NICHOLAS J. SANDERS, AND GREGOR SINGER\*

*Influenza and air pollution each pose significant health risks with global economic consequences. Their shared etiological pathways present a case of compounding health risk via interacting externalities. Using instrumental variables based on changing wind direction, we show that increased levels of contemporaneous pollution increase influenza hospitalizations. We exploit random variation in effectiveness of the influenza vaccine as an additional instrument to show that vaccine protection neutralizes this relationship. Thus, pollution control and vaccination campaigns jointly provide greater returns than those implied by addressing either in isolation. We show the importance of this consideration in addressing observed gaps in influenza incidence by race. (JEL D62, I12, J15, Q51, Q53)*

Influenza (flu) and air pollution are significant public health risks that impact nations around the world. The flu causes an estimated three to five million severe cases per year and nearly half a million deaths (Lambert and Fauci 2010; Iuliano et al. 2018). Air pollution causes 4.5 million annual deaths (Cohen et al. 2017), with annual economic costs estimated to exceed US\$800 billion in the United States alone (Putri et al. 2018; Tschofen, Azevedo, and Muller 2019). While public health policies to address these issues are often considered in isolation, both share common etiological pathways through which they harm human health.

Interactions between the flu and pollution are an illustrative economic case of compounding risk from interacting externalities. Influenza is an infectious disease whereby the actions of one infected individual impose negative externalities on others by increasing risk of infection, while air pollution is a negative externality of economic activity. Our analysis demonstrates that policies to address these distinct externalities have significant interactive effects: the flu vaccine can protect against

\*Graff Zivin: University of California, San Diego (email: [jgraffzivin@ucsd.edu](mailto:jgraffzivin@ucsd.edu)); Neidell: Columbia University (email: [mn2191@cumc.columbia.edu](mailto:mn2191@cumc.columbia.edu)); Sanders: Cornell University (email: [njsanders@cornell.edu](mailto:njsanders@cornell.edu)); Singer: London School of Economics (email: [g.a.singer@lse.ac.uk](mailto:g.a.singer@lse.ac.uk)). Neale Mahoney was coeditor for this article. We thank Douglas Almond, Max Auffhammer, Paulina Oliva, Luisa Osang, Jeffrey Shaman, four anonymous reviewers, and participants at seminars at the University of California Environmental Economics group, University of California San Diego, London School of Economics (LSE), Vanderbilt, AEA-ASSA 2022, UC Santa Barbara Occasional Workshop, and Aarhus for helpful discussions. Singer acknowledges support from the Grantham Research Institute on Climate Change and the Environment at the LSE and the ESRC Centre for Climate Change Economics and Policy (ref. ES/R009708/1). The IRB for access to the HCUP data through the National Bureau of Economic Research (NBER) was approved by the NBER on April 10, 2019, with IRB protocol reference number 19\_135 (FWA number 00003692).

<sup>†</sup>Go to <https://doi.org/10.1257/app.20210500> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

certain harms from air pollution, and reduced levels of air pollution lessen the harmful effects of influenza exposure. Thus, the seemingly disparate policy actions of pollution control and expanded vaccination may jointly provide greater returns than when studied in isolation.

We begin our analysis by exploring the relationship between air pollution and influenza. Exposure to air pollution can affect influenza severity (Jaspers et al. 2005; Lee et al. 2014) and, to a lesser degree, its spread (Chen et al. 2010). We extend the cross-sectional epidemiological literature<sup>1</sup> to establish a causal relationship between air pollution and flu cases. We use patient-level administrative data on inpatient hospitalizations from 2007 to 2017 across 21 US states, which allows us to focus on cases with a definitive influenza diagnosis.<sup>2</sup> We address the identification challenge that vaccine take-up and pollution exposure are often endogenously determined by using a dual instrumental variables approach. We first estimate econometric models with spatial and temporal fixed effects to control for numerous unobservable factors and then build on the pioneering work of Deryugina et al. (2019) by using plausibly exogenous variation in wind directions as an instrument for pollution. We find that higher pollution levels significantly increase flu inpatient hospitalizations; a one standard deviation increase in the monthly air quality index (AQI) (10.9-unit increase in our data) amounts to approximately 35.7 percent additional flu-related inpatient hospitalizations in the United States during influenza season. Compared to the effect of air pollution on all respiratory hospitalizations, our findings suggest influenza accounts for around 18 percent of all air pollution-induced respiratory inpatient hospitalizations.

Next, we explore whether influenza vaccine protection, which we define as a combination of vaccine take-up and effectiveness, moderates the estimated relationship above. As vaccine take-up can be endogenous across both time and location, we instrument for vaccine protection using vaccine effectiveness weighted by influenza susceptibility (in addition to using our instrument for pollution). Effectiveness of the flu vaccine varies from year to year: producers forecast viral strain match months ahead of time, and antigenic drift or shift induces random deviations in realized match quality.<sup>3</sup> This makes the random draw of the viral match orthogonal to unobserved determinants of health, allowing us to identify a causal relationship between the vaccine and health harms from pollution. The orthogonality of vaccine effectiveness also offers an additional test that pollution has a causal effect on flu admissions. If a vaccine designed specifically to protect against the flu diminishes the impact of pollution on influenza hospital admissions, then it must be the case that pollution contributes to influenza hospitalizations. When we include an interaction between air pollution and vaccine protection, we find that the flu vaccine offers significant protection from influenza-related costs of pollution.<sup>4</sup> Vaccine protection levels close

<sup>1</sup> See, for example, Brauer et al. (2002); Wong et al. (2009); Chen et al. (2010); Liang et al. (2014); and the important economic history paper by Clay, Lewis, and Severnini (2018). In a study of the Spanish flu in 1918, Clay, Lewis, and Severnini (2018) show that cities with higher coal-fired power generating capacity saw higher mortality rates, potentially through exposure to higher air pollution.

<sup>2</sup> Estimation based simply on physician encounters is more difficult, as influenza testing is not conducted systematically, and reporting of positive cases is not mandatory for this patient population.

<sup>3</sup> Other papers using similar variation include Ward (2014) and White (2021).

<sup>4</sup> Access to health care may ameliorate impacts of adverse environmental conditions more generally. Mullins and White (2020) show that better access to acute care can help protect against health harms from extreme temperature.

to the average across time in our sample fully neutralize the relationship between pollution and additional flu hospitalizations.

Given the unequal burden of both flu and pollution exposure across society, we also explore results by race and ethnicity. Both of our main findings—that air pollution increases flu hospitalizations and that vaccine protection moderates this relationship—are consistent across these dimensions. Combined with evidence of significant differences in flu incidence and severity by race (e.g., Quinn et al. 2011), our results suggest that the well-established differences in ambient pollution concentrations across racial and ethnic groups (e.g., Banzhaf, Ma, and Timmins 2019; Colmer et al. 2020; Currie, Voorheis, and Walker 2020) serve as an important mechanism driving disparities in influenza outcomes across such groups. Moreover, since flu vaccines protect against some pollution-induced harms, our results imply that the private and external benefits from vaccines are considerably higher in communities disproportionately exposed to poor air quality. Our source of exogenous variation in vulnerability to pollution through vaccine effectiveness also contributes to the literature on the distribution of environmental damages that depends on not only exposure but also vulnerability (Hsiang et al. 2019; Deryugina et al. 2021).

An important feature of our context is that the spreads of influenza and pollution are externalities. As externalities, they justify government intervention in the form of policies, such as increased vaccine take-up and improved air quality.<sup>5</sup> The interaction of the two suggests that the seemingly disparate policy actions of pollution control and vaccination campaigns jointly provide greater returns than those implied by addressing either in isolation. A back-of-the-envelope calculation suggests a 10 percent (3.5 AQI points) reduction in the AQI in a historically ineffective vaccine year (11 percent vaccine take-up adjusted for effectiveness) would avert 16.6 percent of all influenza-associated hospitalizations across the United States. Meanwhile, a 10 percent improvement in vaccine take-up at the average vaccine effectiveness (or, equivalently, a 10 percent improvement in vaccine effectiveness at the average vaccine take-up) in a historically polluted year (38.2 AQI) would avert 34.6 percent of pollution-driven influenza hospitalizations. The optimal mix of these policies will depend on relative costs as well as the spillover benefits each may generate beyond influenza.

The paper proceeds as follows. We begin by describing potential biological mechanisms and our data and present why they are particularly well suited to addressing the question of interacting externalities (Section I). We then discuss our econometric model and describe in detail the various instruments we use to address issues of endogeneity and measurement error (Section II). After we present our main results and explore variations in our model assumptions, we discuss the implications of our findings, in the context of both our analysis and the larger question of social welfare maximization (Section III), before we conclude (Section IV).

<sup>5</sup>A similar logic applies to the more difficult task of improving vaccine effectiveness. In that case, policies are more likely to utilize the standard push-and-pull mechanisms used to overcome the underinvestment problem that arises due to the public good nature of scientific knowledge (Kremer and Williams 2010).

## I. Background and Data

### A. Potential Biological Mechanisms

The primary channel through which air pollution could affect influenza hospitalizations is increasing the severity of influenza. Like smoking (Han et al. 2019), air pollution can impair the respiratory functioning of patients, e.g., by damaging the respiratory epithelium, thereby facilitating the progression of influenza virus beyond the epithelial barrier into the lungs (Diamond, Legarda, and Ryan 2000; Jaspers et al. 2005; Ciencewicki and Jaspers 2007; Rivas-Santiago et al. 2015). Existing medical research finds that exposing in vitro respiratory epithelial cells to air pollution increases the susceptibility and penetration of influenza (Jaspers et al. 2005) and that experimental exposure of mice to air pollution before influenza infections increases morbidity and mortality (Hahon et al. 1985; Lee et al. 2014).

There is also some suggestive evidence that air pollution could affect influenza hospitalizations through modest increases in the spread of influenza. Like humidity and temperature (Lowen et al. 2007; Shaman and Kohn 2009; Shaman et al. 2010; Ijaz et al. 1985; Casanova et al. 2010), air pollution particles could extend the airborne survival of viruses outside the body (Ijaz et al. 1985; Tellier 2009; Chen et al. 2010; Khare and Marr 2015; Lou et al. 2017; Wolkoff 2018) and thus increase the probability of disease transmission.

### B. Data

We combine data from multiple sources on health outcomes, pollution concentrations, vaccine information, and weather variables.

*Inpatient Hospitalizations.*—Our primary health outcome is inpatient hospitalizations for influenza. We use patient-level data on inpatient hospitalizations from the Healthcare Cost and Utilization Project (HCUP 2018b). The HCUP data cover the universe of hospital admissions in the reporting states. While Medicare data, an alternative popular data source, cover the entire United States, the advantage of the HCUP data is that they cover all age groups, not only the elderly. We focus on influenza cases by using patient-level information on diagnosed diseases per International Classification of Diseases (ICD) codes.<sup>6</sup> We limit analysis to data from 2007 to 2017, for which we also have detailed vaccine effectiveness data available. This gives us an unbalanced panel of 21 US states, with an average of 5.5 years of observations per state (see Table A.1 in online Appendix A.1 for details on data availability by state and year). Figure A.1 in the online Appendix shows that the HCUP data is broadly representative of US data by comparing distributions of several sociodemographic variables across counties in our HCUP data with distributions across all US counties.

<sup>6</sup>We exclude patients whose zip code is from a different state than the hospital in which they are treated.

We define our outcome as the count of inpatient admissions per county-year-month where the ICD code indicates influenza.<sup>7</sup> Given the presence of primary and secondary diagnosis codes, we conduct analyses using three possible classifications of flu admissions: cases where (i) the only diagnosis is influenza (most restrictive), (ii) any diagnosis is influenza (least restrictive), and (iii) the primary diagnosis is influenza. The third option reflects a middle ground that we use as our baseline outcome.

We focus on the influenza season, which the US Centers for Disease Control and Prevention (CDC) defines as October to March, and explore results extending the season in online Appendix A.3. Figure 1, panel A shows the seasonality of inpatient hospitalizations in our data, which matches closely with general CDC-reported influenza-like illnesses (see Table A.2 in online Appendix A.1). Based on month of admission and patient zip code, we aggregate hospitalization data to the county-year-month level and assign a zero value to counties in months with no reported influenza admission, conditional on reporting data in the given year.<sup>8</sup> During the influenza season, 54 percent of county-year-months have no reported influenza-related hospital admissions in the HCUP data, and our results are robust to the inclusion or exclusion of zero-valued county-year-months. To compare our main results with the more general effect of air pollution on any respiratory hospitalization (including influenza), we also construct a variable that contains the count of inpatient hospitalizations where the primary diagnosis is any respiratory diagnosis.<sup>9</sup> Finally, for a falsification test, we use primary ICD codes associated with osteoarthritis as an outcome variable, which is unlikely to be affected by air quality and influenza.<sup>10</sup>

*Air Quality.*—As our measure of pollution, we begin with the US Environmental Protection Agency's (EPA's) (2020) AQI at the county-day level, which we aggregate to county-by-year-by-month to match hospitalization outcomes.<sup>11</sup> We focus on the AQI as a summary measure of overall air quality, based on the primary criteria pollutants specified in the Clean Air Act.<sup>12</sup> We do so as the high degree of correlation between several individual pollutants makes it challenging to separately identify the

<sup>7</sup>We use the Clinical Classifications Software from the Agency for Healthcare Research and Quality (AHRQ) to classify relevant influenza ICD codes. These are all five-digit ICD codes grouped under the following three-digit ICD-9-CM codes: 487, 488; and, for the period from October 2015 when the system was changed to ICD-10-CM, the following three-digit ICD-10-CM codes: J09, J10, J11.

<sup>8</sup>Put another way, we only impute zeros for counties and year-months in states that report data in that given year but have zero influenza hospitalizations in a given month. We use the crosswalk from zip codes to counties from the US Department of Housing and Urban Development (Din and Wilson 2020).

<sup>9</sup>These are all five-digit ICD codes grouped under the following two-digit ICD-9-CM codes: 46, 47, 48, 49, 50, 51; and the following two-digit ICD-10-CM codes: J0, J1, J2, J3, J4, J5, J6, J7, J8, J9.

<sup>10</sup>Osteoarthritis consists of all five-digit ICD codes grouped under the following three-digit ICD-9-CM codes: 715, V134; and the following three-digit ICD-10-CM codes: M15, M16, M17, M18, M19. We also show effects for all disease groupings as additional robustness checks in the online Appendix.

<sup>11</sup>The EPA preaggregates data to the daily county level in the case of multiple monitors per county. For missing county-year-months, we take the average value of the adjacent counties in the same month. We winsorize the AQI at the top and bottom 1 percent for the main analysis and show robust results to both data cleaning choices in online Appendix A.3.

<sup>12</sup>The AQI captures pollution from particulate matter (PM<sub>2.5</sub> or PM<sub>10</sub>), sulfur dioxide (SO<sub>2</sub>), carbon monoxide (CO), nitrogen dioxide (NO<sub>2</sub>) and ozone (O<sub>3</sub>). See online Appendix A.1 for descriptive statistics. The EPA provides further details on AQI calculation in EPA (2018).

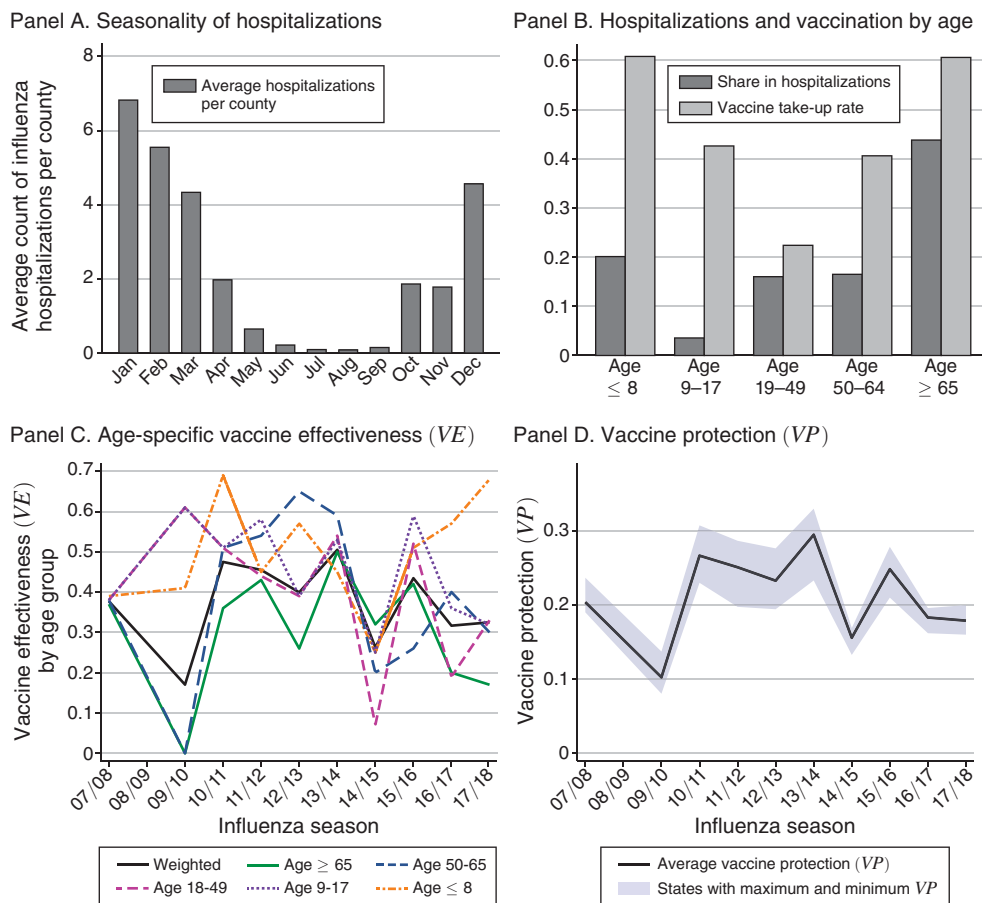


FIGURE 1. DESCRIPTIVE FIGURES ON INFLUENZA INPATIENT HOSPITALIZATIONS AND VACCINE TAKE-UP AND EFFECTIVENESS

Notes: Panel A shows the average count of influenza inpatient hospitalizations per county-month in the HCUP (2018b) data. Panel B shows the age group shares of influenza inpatient admissions as well as age group-specific vaccine take-up, both pooled across states and time. Panel C plots (raw) reported vaccine effectiveness for each age group over influenza seasons (with the exception of 2008/2009, where no data are available). The thick black line plots our weighted measure of overall vaccine effectiveness. Panel D plots vaccine protection averaged across states as the thick line. The bands illustrate the variation within each season across states by plotting the states with the maximum and minimum vaccine protection in each season.

effect of each pollutant independently. We note that most of the “forcing” pollutant that drives variation in the AQI in our setting is PM2.5.

*Weather, Wind Directions, and Inversions.*—To address weather as a confounder, we use monthly weather averages from Xia et al. (2012); Mocko and NASA/GSFC/HSL (2012), including temperature, specific humidity, vertical and horizontal wind speed, and precipitation at the 0.125-by-0.125-degree level, all aggregated up to the county-by-year-by-month level.

To construct our main instrument for pollution, we construct wind direction for a county-year-month by taking the average horizontal ( $u_i$ ) and vertical ( $v_i$ ) wind



components from the monthly raw data and calculating the average angle the wind is blowing from as  $WDIR_i = 180/\pi \arctan 2(-u_i, -v_i)$ .<sup>13</sup>

Temperature inversions can also influence ground-level pollution levels (Arceo, Hanna, and Oliva 2016), which allows us to use inversions as an additional pollution instrument. To calculate inversions, we use daily three-dimensional temperature averages between midnight and 6 AM at each location on each day from the Global Modeling and Assimilation Office (2015), regridded to the 0.25-by-0.25-degree level. We use the difference in temperature between the two pressure levels closest to the surface at each location and average this difference up to the county-day level. We then calculate the share of days with inversions in a county-year-month as the share of days when the difference between the layer further away from the surface and the layer closest to the surface is positive—i.e., the temperature rises with altitude. We calculate the average strength of inversion in a county-year-month as the average difference in temperature between the two altitude levels on the days where inversions are present.

*Vaccine Take-up and Effectiveness.*—We obtain average vaccine take-up rates (VR) by state, season, and age group or racial group from CDC (2008, 2009, 2015, 2020), Lu et al. (2013), and Schiller and Euler (2009). Figure 1, panel B shows that on average, vaccine take-up is highest among those 65 years and older or those 8 years and younger. Figure A.4a in online Appendix A.1 shows temporal variation in vaccine take-up rates by age group, and Figure A.4b by race. Figure A.4c shows spatial variation by taking a cross section of vaccine take-up rates among those 65 years and older across different states in a given influenza season, in this case 2009/2010. The figures illustrate that the variation in vaccine take-up is larger across age groups than across racial groups, time, or space.

We obtain measures of vaccine effectiveness by influenza season and age group,  $VE^{raw}$ , from the studies underlying CDC estimates (CDC 2019), available beginning in the 2007/2008 season (Belongia et al. 2011; Griffin et al. 2011; Treanor et al. 2012; Ohmit et al. 2014; McLean et al. 2015; Gaglani et al. 2016; Zimmerman et al. 2016; Jackson et al. 2017; Flannery et al. 2019; Rolfes et al. 2019; Flannery et al. 2020) with the exception of the 2008/2009 season.<sup>14</sup> These studies measure vaccine effectiveness as the vaccination-induced percentage reduction in the odds of testing positive for influenza conditional on having influenza-like symptoms. One can interpret vaccine effectiveness as the approximate share of vaccinated people who do not test positive but would have absent the vaccine.<sup>15</sup>

Figure 1, panel C plots age-specific vaccine effectiveness against influenza season, showing variation across both seasons and age groups. Across seasons, the match between circulating viral strains and the vaccines based on forecasts is imperfect and varies due to antigenic drift. Within a season, the match can be of different quality for different age groups due to “original antigenic sin” (Francis 1960); the

<sup>13</sup>We calculate wind speed for our control variables as  $WSPEED_i = \sqrt{u_i^2 + v_i^2}$ .

<sup>14</sup>The CDC measures vaccine effectiveness across influenza seasons rather than calendar years, as seasons overlap calendar years (e.g., October–December for year  $y$  and January–March for year  $y + 1$ ).

<sup>15</sup>The odds ratio is approximately the relative risk due to a small number of influenza positive cases (Zhang and Kai 1998).

first influenza strain to which the immune system is exposed imprints immunological memory with that specific strain, such that different generations with different antigenic imprints respond differently to new vaccines and strains within years.

*Constructing Vaccine Protection.*—The share of people protected by the vaccine in each season and state is a combination of take-up rate,  $VR$ , and age group-weighted vaccine effectiveness,  $VE$ . As an example, for a group with homogeneous effects from exposure, if 50 percent of people are vaccinated but the vaccine is only 30 percent effective, the effective vaccine protection ( $VP$ ) is the same as when only 30 percent of people are vaccinated but the vaccine is 50 percent effective. For groups with heterogeneous vulnerability, aggregate hospitalizations also depend on whether those individuals that are more vulnerable than others have a higher take-up rate or vaccine effectiveness. An 80-year-old without a vaccine, for example, is much more likely to be hospitalized with influenza than a 30-year-old without a vaccine. Figure 1, panel B shows hospitalization incidence is highest for two age groups: 65 years and older and 8 years and younger.<sup>16</sup> To construct a population-level measure of vaccine protection that accounts for such differences in vulnerability, we weight age-specific vaccine take-up rates and vaccine effectiveness by influenza hospitalization shares of each age group:

$$(1) \quad VP_{cs} = \frac{1}{\sum_a (\overline{HS}_a)} \sum_a VE_{sa}^{raw} \times VR_{csa} \times \overline{HS}_a,$$

where  $c$  denotes counties ( $VR_{csa}$  varies at the state level, but we index by counties for simpler notation in the following sections),  $s$  denotes influenza seasons, and  $a$  denotes age groups. Hospitalization weights  $\overline{HS}_a$  are a simple average across influenza seasons  $s$ , i.e.,  $\overline{HS}_a = \frac{1}{S} \sum_s HS_{sa}$ , and the first term  $1/\sum_a (\overline{HS}_a)$  ensures that the age weights sum to one, such that overall hospitalizations do not affect our values of  $VP_{cs}$ . We plot  $VP_{cs}$  averaged across states in Figure 1, panel D, along with the  $VP_{cs}$  of the state with the highest and lowest  $VP_{cs}$  in each influenza season. The minimum of  $VP_{cs}$  is 0.08, and the maximum is 0.33.

Since  $VP$  is constant within the season, vaccination rates  $VR_{csa}$  that differ across states solely drive cross-sectional spatial variation in  $VP_{cs}$ . The sources of temporal variation in  $VP_{cs}$  are both vaccination rates,  $VR_{csa}$ , and vaccine effectiveness,  $VE_{sa}^{raw}$ , which vary across influenza seasons. Equation (1) shows that a 10 percent increase in  $VP_{cs}$  can be the result of either a 10 percent increase in vaccine rates in all age groups or a 10 percent increase in vaccine effectiveness in all age groups (or some combination of both effects). For our analysis of heterogeneity across different age groups, we only use vaccination rates and vaccine effectiveness for the relevant age groups in constructing  $VP_{cs}$ . For heterogeneity analysis across different racial groups, we use our overall measure of vaccine protection scaled by the ratio of race-specific take-up in a season to overall vaccine take-up in a season.

<sup>16</sup>We construct groups with these age cutoffs because they coincide with the common age cutoffs in vaccine effectiveness studies.



*Mortality and Emergency Department Visits.*—Although our primary focus is on inpatient hospitalizations, we also extend our analysis to consider influenza-related emergency department (ED) visits and mortality. Data on visits to EDs are from HCUP (2018a) and have overlapping spatial coverage with our main inpatient data. Individual-level mortality data from NCHS (2019) cover every county in the United States and includes deaths that happen inside or outside of hospitals. For both ED visits and mortality, we count every hospitalization or death with influenza as primary cause as above and aggregate to the county-by-year-by-month level.

*Sociodemographics.*—We use employment counts at the county-by-year-by-month level from the Bureau of Labor Statistics (2021) as an additional control in robustness checks. Our analysis of policy implications utilizes county population data by race from the 2010 US census (US Census Bureau 2020) and county median income data from Chetty et al. (2018).

## II. Empirical Strategy

Given the nature of our outcome variables, we estimate count models as our primary specification, though we also estimate linear models as a specification check. We estimate the relationship between the count of influenza-related inpatient hospitalizations  $H_{cym}$  and the lagged air quality index  $AQI_{cym-1}$  at the county  $c$ –by–year  $y$ –by–calendar month  $m$  level using the following conditional exponential mean function (consistent with a Poisson count data model):

$$(2) \quad E[H_{cym} | AQI_{cym-1}, \mathbf{X}_{cym}, \gamma_{csy}, \mu_{ym}] \\ = \exp(\beta AQI_{cym-1} + \mathbf{X}'_{cym} \boldsymbol{\delta}_1 + \mathbf{X}'_{cym-1} \boldsymbol{\delta}_2 + \gamma_{csy} + \mu_{ym}).$$

We lag the AQI one month to capture exposure to air pollution before hospital admission, and we control for a wide variety of both regional and temporal factors. Our preferred specification includes county-by-season-by-year ( $\gamma_{csy}$ ) and year-by-month fixed effects ( $\mu_{ym}$ ). Since each influenza season  $s$  spans October through March and overlaps calendar years  $y$  and  $y + 1$ , the county-by-season-by-year fixed effects ( $\gamma_{csy}$ ) are tantamount to county-by-quarter-by-year fixed effects.<sup>17</sup> While county-by-quarter-by-year fixed effects capture the bulk of climatic differences across counties, we also include contemporaneous weather controls  $\mathbf{X}_{cym}$  and lagged weather controls  $\mathbf{X}_{cym-1}$  to address the links between both influenza and weather (temperature and humidity can influence influenza transmission rates) and weather and pollution (different climatic conditions can lead to different levels of air quality) within county-quarter-years.<sup>18</sup> Note that our outcome variable is the count of

<sup>17</sup>The county-by-season-by-year fixed effects ( $\gamma_{csy}$ ) are equivalent to including county-by-year and county-by-season fixed effects separately.

<sup>18</sup>This includes information on temperature, specific humidity, precipitation, and wind speed. Temperature and humidity have been shown to affect both virus survival (Lowen et al. 2007; Shaman and Kohn 2009; Shaman et al. 2010; Casanova et al. 2010; Harper 1961) and air pollution (Ijaz et al. 1985; Lou et al. 2017; Greenburg et al. 1967). In our baseline model, we include five quintile bins for temperature (C), five quintile bins of specific humidity, and linear terms for precipitation and wind speed, all of which include contemporaneous and lagged versions.

influenza hospitalizations, but inclusion of our fixed effects in equation (2) ensures that our estimates are equivalent to recovering the effect on hospitalization rates per capita.<sup>19</sup>

County-by-season-by-year (or county-by-quarter-by-year) effects  $\gamma_{c, sy}$  control for differences in unobserved confounders that influence pollution exposure and health outcomes across counties separately for every quarter-by-year, such as demographics, socioeconomic factors, or health-care access and protocols. This also addresses a possible concern due to potential variation in random diagnostic influenza testing in hospitals that could mask true influenza rates. Our fixed effects absorb potential bias from discrepancy between actual and observed hospitalizations as long as the ratio between them is constant within county-quarter-years and/or year-months.<sup>20</sup> Year-by-month fixed effects control for seasonality and general monthly trends within each year in both influenza and pollution. For example, two common lung irritants included in the AQI, particulate matter and carbon monoxide, peak in winter months much like influenza admissions; year-by-month fixed effects capture such seasonality. In robustness checks, we examine models using alternative fixed effects specifications.

Given the included fixed effects, two remaining threats to identification are unobserved confounding within each county-by-quarter-by-year cell and measurement error in pollution assignment. For example, increased economic activity and interaction between people at the local level could drive both air pollution and influenza infections. We control for lagged employment at the county-by-year-by-month level in our regressions as one approach, but this may not fully capture this relationship. Measurement error may arise because of the sparse nature of pollution monitors. To address these issues, our main strategy is to employ instrumental variables for air quality.

### A. Instrumenting for Air Quality

We take an approach conceptually akin to that of Deryugina et al. (2019) by exploiting changes in wind direction as an instrument for the AQI. The idea behind the instrument is that wind blowing from a particular direction moves around the pollution internal to the county but also brings in external pollution. Online Appendix Figure A.5 provides an illustration using Suffolk County, which contains the city of Boston, as an example. The polar plot indicates that monthly AQI is much higher when prevailing winds are blowing from the southwest, which is the direction of major polluting sources along the Eastern Seaboard. The AQI is, however, much

<sup>19</sup>Influenza hospitalizations rates  $H_{cym}^{rates}$  and counts  $H_{cym}$  relate in the following way:  $H_{cym}^{rates} = H_{cym}/Pop_{c, sy}$ , where  $Pop_{c, sy}$  is county population in each county-season-year cell. We can multiply both sides of equation (2) by  $\exp(\log(1/Pop_{c, sy}))$  such that our estimation recovers the effect on  $H_{cym}^{rates}$  as dependent variable and the fixed effects absorb  $\exp(\log(1/Pop_{c, sy}))$ .

<sup>20</sup>Suppose actual (unobserved) influenza hospitalizations  $H_{cym}^{actual}$  and measured diagnosed influenza hospitalizations  $H_{cym}$  relate in the following way:  $H_{cym}^{actual} = H_{cym} \times R_{c, sy} \times R_{ym}$ , where  $R_{c, sy} \times R_{ym}$  captures arbitrary discrepancy between actual and observed hospitalizations. If we insert this relationship in equation (2), we can multiply both sides by  $\exp(\log(R_{c, sy}) + \log(R_{ym}))$  such that our estimation recovers the effect on the unobserved  $H_{cym}^{actual}$  as dependent variable and the fixed effects absorb  $\exp(\log(R_{c, sy}) + \log(R_{ym}))$ .

lower when prevailing winds are blowing from the east, which brings in cleaner air from the ocean. We are not using these values directly as our instrument for air pollution, as individuals may sort based on prevailing wind patterns. Instead, we use monthly changes in wind directions as our source of variation that shifts pollution, net of average pollution in the county. For example, if prevailing winds change from southwest to east in Boston, our instrument would shift pollution down for that particular month, net of average pollution levels in Boston in that year-quarter, conditional on our fixed effects and controls. The identifying assumption is that, conditional on our weather controls and fixed effects, wind direction affects influenza hospitalizations only through its effect on the AQI but does not have a direct effect on hospitalizations. Those exogenously determined changes in wind direction (conditional on wind speed, other weather controls, and the various fixed effects) result in changes in pollution levels in a neighborhood that are likely to be uncorrelated with local determinants of pollution.

While we borrow the premise of this design from Deryugina et al. (2019), we modify the precise construction of the instruments. Specifically, Deryugina et al. (2019) construct instruments ( $\mathbf{Z}_i^D$ ) by using dummy variables for wind direction bins  $WDIR_i^q$  (e.g.,  $WDIR_i^{NW}$  for when wind is blowing from the northwest for observation  $i$  belonging to a particular county in a particular point in time) interacted with geographical region-level indicators  $G_c$ :  $\mathbf{Z}_i^D = \sum_c \sum_q WDIR_i^q \times G_c$ . One challenge in constructing this set of instruments is the choice of geographical granularity for  $G_c$ . On the one hand, if  $G_c$  are large regions including multiple counties, a particular wind direction requires that pollution shifts in the same direction and to the same degree for all counties in the same group  $G_c$ . Counties just north or just south of an urban center, however, are likely to receive the pollution shock when wind blows from the opposite direction rather than from the same direction. Similarly, a county south of a large urban center and a county south of a small urban center should receive a pollution shock when wind is blowing from the north, but the size of the pollution shock likely differs. On the other hand, if  $G_c$  are small entities—e.g., counties themselves—then each county is allowed to have different pollution shocks in different sizes from different wind directions, but the set of instruments grows larger than the number of panels or counties  $N_c$ . This can lead to computational difficulties and inefficient standard errors.<sup>21</sup> Deryugina et al. (2019) balance this trade-off by selecting the granularity of  $G_c$  based on a k-means cluster algorithm, which generates groups that include nine counties on average.

We instead solve this trade-off by using a different approach that allows full flexibility in how wind directions shift pollution in different counties (i.e.,  $G_c$  at the county level) while dramatically reducing the number of instruments as well. Instead of interacting wind direction bins with county indicators, we transform the values in the wind direction dummies to capture both the sign and size of pollution shocks from changes in wind direction for each county. We do this in two steps. First, we create a new variable,  $A\tilde{Q}I^{qc}$ , which is pollution in county  $c$  averaged over

<sup>21</sup> Optimal (two-step) generalized method of moments (GMM) with a clustered weighting matrix at the county level is infeasible, for example, because the number of instruments is larger than the number of clusters.

the entire sample when wind is blowing from direction  $q$  in county  $c$ , demeaned by the average pollution level in county  $c$ :

$$(3) \quad A\tilde{QI}^{q_c} = \frac{1}{\sum_{i \in q_c} \sum_{i \in q_c} A Q I_i^{q_c}} - \frac{1}{\sum_{i \in c} \sum_{i \in c} A Q I_i}.$$

We then use  $A\tilde{QI}^{q_c}$  to generate a set of instruments  $\mathbf{Z}_i^q$ , where each instrument corresponds to a particular wind direction (e.g.,  $Z_i^{NW}$ ) and the values of  $\mathbf{Z}_i^q$  are populated by  $A\tilde{QI}^{q_c}$  if a particular observation  $i$  belongs to county  $c$  and the wind in this particular year-month in this county is blowing from  $q$ :

$$(4) \quad \mathbf{Z}_i^q = \begin{cases} A\tilde{QI}^{q_c}, & \text{if } WDIR_i^q = q \text{ and } i \in c; \\ 0, & \text{otherwise.} \end{cases}$$

This generates  $N_q$  instruments instead of  $N_q \times N_c$  instruments.  $\mathbf{Z}_i^q$  also addresses the two restrictions that arise when pooling multiple counties into groups. First, a single coefficient on a particular wind direction bin (e.g., the coefficient for  $Z_i^{NW}$ ) accounts for different signs of pollution shocks for different counties from the same wind direction. For example, a county southeast of a major urban center is likely to have a positive value in  $Z_i^{NW}$ , whereas a county northwest of the major urban center is likely to have a negative value in  $Z_i^{NW}$ . Therefore, the coefficient for  $Z_i^{NW}$  can shift pollution for the two counties into different directions. Second, a single coefficient on a particular wind direction bin also accounts for different sizes of pollution shocks. For example, a county southeast of a large urban center may experience larger pollution shocks when wind blows from the northwest than a county southeast of a small urban center. Since the average size of pollution shocks is captured in  $Z_i^{NW}$ , the same coefficient on  $Z_i^{NW}$  can shift pollution to a different extent in different counties.

We design the instruments  $\mathbf{Z}_i^q$  to capture pollution shocks that occur from changes in wind direction. Since we use wind-induced pollution shocks averaged across the entire sample when constructing  $\mathbf{Z}_i^q$ , we do not capture individual events that generate pollution shocks that only occur in a particular year-month in a particular county, which could be problematic since they may also correlate with influenza cases.<sup>22</sup> We also control for changes in weather that might affect influenza hospitalizations directly and correlate with changes in wind direction, such as temperature, humidity, precipitation, or wind speed. Finally, since we use a one-month lagged AQI as our variable of interest, we use a one-month lagged wind direction instrument to form our moment conditions.

For our baseline model, we use the four quadrants as wind direction bins but have also performed robustness checks with alternative numbers of wind direction

<sup>22</sup>Identification does not use prevailing wind direction, which would not change across time and, as Deryugina et al. (2019) note, could lead to sorting or strategic placement of pollution monitors. Instead, our instrument uses month-to-month changes in wind patterns in a given county-by-year-by-quarter cell, which should not affect sorting or monitor placement.

bins. We estimate our instrumented model with a Poisson GMM-IV procedure that accounts for fixed effects through quasi-mean differencing, and we construct moment conditions with our set of instruments. Note that the noninstrumented Poisson GMM estimates are numerically equivalent to a Poisson pseudo-maximum likelihood (PPML) estimator.<sup>23</sup> We cluster standard errors at the county level to allow for arbitrary heteroskedasticity and serial correlation. For our linear specification, we use the corresponding linear GMM-IV procedure that is numerically equivalent to standard linear GMM optimization. We provide econometric details in online Appendix A.2.

As an expansion, we include further instruments for AQI based on thermal inversions (Arceo, Hanna, and Oliva 2016). Typically, air is colder the farther it is from the earth's surface. Thermal inversions appear when a warm air layer moves above a cold air layer, reducing air cycling and generating stagnant air conditions. While inversions do not directly affect health (conditional on temperature), they trap pollutants closer to the ground, leading to increases in pollution concentrations.<sup>24</sup> We use the share of days with inversions and the average strength of inversions at the county-year-month level. We then interact both variables with a scaling variable that is the average county AQI across the entire sample. This allows inversions in more pollution-intensive regions (e.g., large urban centers) to shift pollution more than in less pollution-intensive regions (e.g., rural counties).

While our Poisson GMM-IV fixed effects estimation does not have an explicit first-stage regression as in two-stage least squares estimations, we can approximate a first stage by running a linear regression of AQI on our instruments and controls. Table A.3 in online Appendix A.3 shows that our wind instruments shift pollution with a Kleibergen-Paap  $F$ -stat of 176.8 (column 1).<sup>25</sup> Inversions also shift pollution, however; the Kleibergen-Paap  $F$ -stat is lower at 8.6 when including inversions alone (column 4) and at 91 when including wind direction and inversion instruments simultaneously (column 7).<sup>26</sup> For this reason, our preferred specification relies solely on the instruments based on wind direction, though we also show results with both sets of instruments.

### B. Vaccines

To estimate the impact of vaccine protection ( $VP_{cs}$ ) on the pollution-hospitalization relationship, we modify equation (1) to include an interaction term

<sup>23</sup>We show the PPML (Correia, Guimarães, and Zylkin 2019) estimates in the online Appendix. The PPML point estimates are consistent as long as the conditional mean is correctly specified, irrespective of the distribution of the outcome or errors (Gourieroux et al. 1984). The PPML estimator performs well with a large number of zeros and over- or underdispersion in the data (Silva and Tenreiro 2006, 2011).

<sup>24</sup>We use inversions between midnight and 6 AM to limit potential confounding through behavioral responses.

<sup>25</sup>Note that all wind direction bins have a positive coefficient, because the values of the instrument are negative when a particular wind direction tends to blow in clean air for a particular county. The coefficients should converge to one as the sample size grows, through either the number of years or the number of counties. Table A.4 shows this pattern in a Monte Carlo simulation of the approximated first-stage regression.

<sup>26</sup>Note that the sum of the two coefficients, the coefficient on the interaction between share of inversion days with the county average AQI ( $\overline{AQI}$ ) and the coefficient on share of inversion days, is positive at the average of  $\overline{AQI}$  (34.7), and the same holds for the strength of inversions.

$AQI_{cym-1} \times VP_{cs}$ , noting that the base effect  $VP_{cs}$  is absorbed in the fixed effects  $\gamma_{csy}$ :

$$(5) \quad E[H_{cym} | AQI_{cym-1}, VP_{cs}, \mathbf{X}_{cym}, \gamma_{csy}, \mu_{ym}] \\ = \exp(\beta_1 AQI_{cym-1} + \beta_2 (AQI_{cym-1} \times VP_{cs}) + \mathbf{X}'_{cym} \delta_1 \\ + \mathbf{X}'_{cym-1} \delta_2 + \gamma_{csy} + \mu_{ym}).$$

Several econometric challenges exist in evaluating how the influenza vaccine alters the effect of pollution on influenza. Recall that vaccine protection,  $VP_{cs}$ , is a composite measure of vaccine take-up and effectiveness. Individuals may reduce avoidance behavior if vaccinated, or be more likely to get the vaccine in seasons with more reported influenza cases, both of which attenuate the raw effect of the vaccine. Selection bias in vaccine take-up may also pose a problem if the most susceptible or most cautious are more likely to seek out vaccines. To address these issues, we instrument for potentially endogenous vaccine protection ( $VP_{cs}$ ) using exogenous vaccine effectiveness ( $VE_s$ ). Our identifying variation exploits the natural variation in vaccine effectiveness, determined by the random variations in the quality of the match between the influenza vaccine and the viral strain in circulation.<sup>27</sup> Note that at the time of vaccination, which is usually early in the influenza season, it is not yet known how effective the vaccine will turn out over the course of the season. Therefore, vaccine take-up should generally not be affected by vaccine effectiveness. We confirm this empirically by regressing take-up on effectiveness separately for our five age groups and find no statistically significant association in any of the five regressions, with estimated implied elasticities close to zero (see online Appendix Table A.5).

Effectiveness based on antigenic drift is, in principle, orthogonal to unobserved determinants of health in a given year. This provides insights into how vaccines affect the pollution-induced spread of influenza and provides a test of the causal effects of pollution on influenza. If vaccines moderate the effect of pollution on influenza, it must be that pollution causally relates to influenza hospitalizations, though we cannot distinguish between whether the vaccine is (i) reducing the probability that any pollution-harmed individual is exposed to the flu due to external benefits from vaccination of others or (ii) changing the probability that a pollution-harmed individual contracts a severe case of flu when exposed.

To generate an overall measure of vaccine effectiveness ( $VE_s$ ) to instrument for  $VP_{cs}$ , we construct a weighted average of time-varying age-specific raw vaccine effectiveness ( $VE_{sa}^{raw}$ , which Figure 1, panel C shows). The weights for age groups are time invariant and capture the age groups where vaccine effectiveness matters relatively more: those with a greater tendency of hospitalization and those with

<sup>27</sup> See also Ward (2014) and White (2021), who, however, calculate vaccine effectiveness based on the names of the viral strains in the vaccine and in circulation, which, in contrast to our measure, do not take into account variations in vaccine effectiveness across age groups and imperfectly map into clinical measures of effectiveness.



higher vaccine take-up rates. Figure 1, panel B shows these weights and that both hospitalization incidence and vaccination rates are highest for those 65 years and older and those 8 years and younger, the two most vulnerable groups in our sample. Our measure of vaccine effectiveness is

$$(6) \quad VE_s = \frac{1}{\sum_a (\overline{VR}_a \times \overline{HS}_a)} \sum_a VE_{sa}^{raw} \times \overline{VR}_a \times \overline{HS}_a,$$

where vaccine take-up rate weights  $\overline{VR}_a$  and hospitalization shares  $\overline{HS}_a$  are simple averages across influenza seasons  $s$ —e.g.,  $\overline{VR}_a = \frac{1}{S} \sum_s VR_{sa}$ —and the first term  $1/\sum_a (\overline{VR}_a \times \overline{HS}_a)$  ensures that the age weights sum to one such that overall vaccine take-up or hospitalizations do not affect our values of vaccine effectiveness. As we use time-averaged hospitalization shares and vaccination rates, vaccine effectiveness is the only source of temporal variation in our instrument.<sup>28</sup> Figure 1, panel C shows that our final measure of weighted vaccine effectiveness ranges between 0.17 and 0.51 during our study period.

By defining vaccine protection as a combination of vaccine effectiveness and vaccine take-up, we interpret  $\beta_2$  as a change in either component, suggesting that policy can focus on either measure. This helps maintain a direct policy implication of our results—while random variation in vaccine effectiveness provides a compelling identification strategy, policy efforts to improve it are met with limited success. Vaccine take-up rates, however, may be more amenable to policy intervention through efforts to reduce the costs of obtaining a vaccine or promote its benefits. With this policy lens in mind, we discuss changes in  $\beta_2$  as the effect of a relative increase in vaccine take-up rates.

To estimate equation (5), we use the same Poisson GMM-IV fixed effects estimator as for equation (2), with wind direction instruments for the AQI. The moment conditions for our interaction term  $AQI \times VP$  use the interaction of wind direction instruments with our  $VE$  instrument. Table A.3 in online Appendix A.3 shows that our wind instruments interacted with  $VE$  shift the interaction term with a Kleibergen-Paap  $F$ -stat of 35.3 (column 3).

### III. Results and Discussion

#### A. Influenza Hospitalizations

Table 1 shows estimates from our Poisson GMM estimations. Coefficients represent the AQI semielasticity of the count of inpatient hospitalizations with primary-diagnosis influenza within a county-year-month or an approximate percentage change in inpatient counts per unit of AQI when estimates are sufficiently small. Estimates from column 1 in panel A correspond to equation (2), without using any instruments, and

<sup>28</sup> A potential threat to the exclusion criteria for our instrument occurs if shocks that increase the spread of influenza (e.g., a sporting event associated with a local team, as in Stoecker, Sanders, and Barreca (2016)) also increase influenza mutation rates and thus weaken vaccine effectiveness. This concern is likely limited, as recent research suggests that random mutations during vaccine production, not from virus spread itself, drive mismatch of vaccines and strains in circulation. For discussion of this research, see Cohen (2017).

TABLE 1—THE EFFECT OF AIR POLLUTION ON SEVERE INFLUENZA CASES

	Influenza is primary ICD code		Influenza is any ICD code		Influenza is only ICD code	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Poisson GMM</i>						
<i>AQI</i>	0.0076 (0.0024)	0.034 (0.0076)	0.0082 (0.0024)	0.031 (0.0070)	0.014 (0.0058)	0.037 (0.020)
<i>AQI</i> × <i>VP</i>		−0.14 (0.036)		−0.12 (0.032)		−0.13 (0.10)
<i>Panel B. Poisson GMM-IV</i>						
<i>AQI</i>	0.028 (0.0074)	0.11 (0.026)	0.021 (0.0069)	0.088 (0.024)	0.043 (0.017)	0.11 (0.049)
<i>AQI</i> × <i>VP</i>		−0.53 (0.16)		−0.41 (0.14)		−0.49 (0.32)
Observations	17,668	17,668	20,013	20,013	3,954	3,954
Mean of outcome	6.04	6.04	11.05	11.05	0.81	0.81
Mean of <i>AQI</i>	35.27	35.27	35.06	35.06	38.07	38.07
Mean of <i>VP</i>	—	0.21	—	0.21	—	0.2
Mean of <i>VE</i>	—	0.36	—	0.36	—	0.35

*Notes:* The dependent variable in columns 1–2 is the count of inpatient hospital admissions with influenza as the primary diagnosis within a county-year-month. The dependent variable in columns 3–4 is the count of inpatient hospital admissions with influenza as any (primary or secondary) diagnosis within a county-year-month. The dependent variable in columns 5–6 is the count of inpatient hospital admissions with influenza as the only diagnosis within a county-year-month. We limit analysis to the influenza-intensive months of October through March, and our sample spans 2007–2017 with the exception of October 2008 to March 2009, where vaccine effectiveness data are not available. Vaccine protection (*VP*) is weighted by hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson GMM estimation with county-by-season-by-year fixed effects and year-by-month dummies as well as weather controls. Weather controls consist of five bins of temperature quintiles, five bins of specific humidity quintiles, and linear terms for precipitation and wind speed. All weather variables are based on county-year-month averages. The air quality index (*AQI*) is lagged one month, and a higher *AQI* means worse air quality. The columns in panel B use our instruments based on wind direction instead of the *AQI* to generate moment conditions, and even-numbered columns additionally use our *VE* instrument instead of *VP* to form moment conditions. The number of included observations can vary across different outcomes due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level.

imply a 1-unit increase in the monthly *AQI* associates with a 0.76 percent increase in influenza inpatient admissions. Column 1 in panel B shows that the estimate is larger when using instruments for the *AQI* based on wind direction. Given that the noninstrumented estimates contain county-by-quarter-by-year fixed effects, they likely control for many potential sources of confounding, such as residential sorting or economic activity. Much of the remaining bias in the noninstrumented estimates is likely due to measurement error. If this error is classical, estimates will be biased to the null, and the IV approach will generate larger (absolute) estimates. Our larger IV estimates are consistent with this, as well as the patterns found in Deryugina et al. (2019).<sup>29</sup> Specifically, our IV approach finds that a 1-unit increase in the monthly *AQI* results in a 2.8 percent increase in influenza inpatient admissions. To put this estimate into a national context, a one standard deviation increase in *AQI* (10.9-unit

<sup>29</sup>The *p*-value of Hansen's *J*-statistic of overidentifying restrictions in column 1 in panel A is 0.53, so we cannot reject validity of the model.

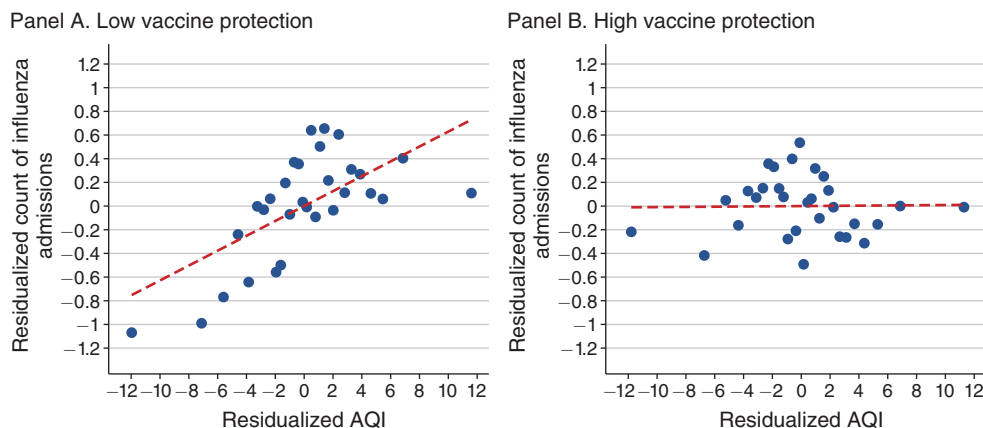


FIGURE 2. AIR QUALITY AND VACCINE PROTECTION

*Notes:* Panels A and B show binned scatterplots with 30 bins and a linear regression on the underlying data. Each shows the correlation net of county-by-season-by-year and month fixed effects as well as weather controls, where the vertical axis shows the residuals from a Poisson regression and the horizontal axis the residuals from a linear regression without instruments. The panels show the relationship for below- (A) and above-median (B) vaccine protection in the sample.

increase in our data) amounts to approximately 27,182 (35.7 percent) additional inpatient hospitalizations for a 6-month influenza season in the United States.<sup>30</sup>

To explore the moderating role of the influenza vaccine, Figure 2 shows the regression-adjusted relationship between AQI and influenza admissions separately in a sample with low vaccine protection in panel A and high vaccine protection in panel B. We determine each group using a median vaccine protection (0.21) sample split. The relationship between air quality and admissions rates is positively sloped in panel A, indicating that the AQI affects flu admissions when the vaccine is a bad strain match and/or vaccine take-up is low. When vaccine protection is high, however, this relationship flattens almost completely, as panel B shows, suggesting that an effective vaccine with sufficient take-up nullifies the relationship between pollution and the flu. This does not imply that a high vaccine protection eliminates all influenza hospitalizations or all pollution-related respiratory hospitalizations. Rather, sufficiently high vaccine effectiveness and take-up eliminate those flu hospitalizations directly attributable to the negative shock of pollution.

To test for the moderating role of vaccine protection, we present estimates of equation (5) using our Poisson GMM framework in Table 1. Column 2 in panel A shows the estimates without using instruments, and column 2 in panel B uses

<sup>30</sup>We use the 10.9-unit increase and the coefficient 0.028 for the relative increase  $\exp(0.028 \times 10.9) - 1 = 0.3569$ , and multiply it by the average inpatient admissions per county-year-month (4.04), the total number of US county equivalents according to the US Census Bureau (3142) (United States Census Bureau 2018) and by the 6 months within a influenza season. Note that we are using average admissions across our pre-estimation sample of summary statistics from Table A.2 (4.04), which is lower than the average reported in the estimation sample in Table 1 (6.04), since a count model drops counties with zero-valued outcomes within the level of the fixed effect. This only counts cases with primary-diagnosis influenza, making this estimate of absolute numbers a lower bound. Using hospitalization with any influenza diagnosis (column 3) doubles the additional predicted cases because the base of hospital admissions is much larger.

our instruments based on wind direction for the AQI and our vaccine effectiveness instrument (*VE*) interacted with the wind direction instruments for the interaction term of AQI and vaccine protection (*VP*). The instrumented estimates are larger than the noninstrumented estimates by around the same factor as for the noninteracted results in column 1 in panel A and B. Vaccine protection substantially moderates pollution-driven influenza cases. Our negative interaction coefficient in column 2 in panel B implies that a vaccine protection of 21 percent, which coincides with the average vaccine protection in our sample (the maximum is 33 percent), nullifies the link between air pollution and influenza hospitalizations. This supports prior evidence of thresholds in influenza vaccination where the positive external benefits are large enough to almost eliminate influenza spread even at incomplete vaccination take-up and effectiveness (Boulier, Datta, and Goldfarb 2007; Ward 2014). In seasons with poor viral match of the vaccine (see Figure 1, panel C), vaccine protection is substantially lower (see Figure 1, panel D). To compensate for a drop in vaccine effectiveness from the median (0.39) to the twenty-fifth percentile (0.32), vaccine take-up would need to increase by 18 percent across all age groups. Table A.6 in online Appendix A.3 provides reduced-form results where we include vaccine effectiveness directly instead of instrumenting for vaccine protection.

In our baseline specifications in columns 1 and 2, we include only cases where the primary diagnosis is influenza, thus ignoring occurrences of influenza in secondary diagnoses. This likely misses some influenza-related hospitalizations but is arguably more robust to overcounting cases that might arise by including patients who suffer from different health conditions triggered by air pollution (e.g., asthma) and then happen to be tested for influenza upon hospital admission due to health protocols. To show robustness to different counting strategies, columns 3 and 4 repeat our analysis counting patients that have any (primary or secondary) influenza diagnosis. This yields an average number of influenza admissions per county-year-month in our estimation sample that is roughly double (11.05) compared to our baseline approach (6.04). The estimated coefficients, which again reflect semielasticities, are close to baseline results both for the level effect of AQI as well as the interaction with vaccine effectiveness. In columns 5 and 6, we use a more restrictive condition by counting hospital admissions where the only diagnosis is influenza. This reduces the average count of admissions per county-year-month to 0.81 (the majority of influenza hospital admissions have further influenza-induced complications—e.g., pneumonia). The estimated coefficients are again comparable to our baseline estimates, though with larger standard errors given the considerable drop in sample size due to more cells with zero counts.

Table 2 explores heterogeneity by age and race using our Poisson GMM-IV specifications (we show noninstrumented results in Table A.7 in online Appendix A.3).<sup>31</sup> Columns 1 through 6 in panel A show results for three distinct age groups: up to age 8, ages 9 through 64, and ages of at least 65 years, where the first and last reflect

<sup>31</sup> For our regressions with age-specific outcomes in Table 2, we only use the vaccine take-up rate and raw vaccine effectiveness data of the corresponding age groups for constructing our overall measure of vaccine protection (*VP*) and vaccine effectiveness (*VE*). We show means of *VP* and *VE* for each regression at the bottom of the table. We note that vaccines have private and external benefits, so vaccine take-up of any one group generates positive spillovers to other groups.

TABLE 2—HETEROGENEITY BY AGE AND RACE

<i>Panel A. By age (Poisson GMM-IV)</i>						
	≤ 8 years		9–64 years		≥ 65 years	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>AQI</i>	0.034 (0.0093)	0.13 (0.051)	0.032 (0.0080)	−0.039 (0.054)	0.0050 (0.013)	0.037 (0.014)
<i>AQI</i> × <i>VP</i>		−0.34 (0.16)		0.45 (0.34)		−0.33 (0.15)
Observations	10,593	10,593	13,984	13,984	13,619	13,619
Mean of outcome	1.89	1.89	2.76	2.76	3.51	3.51
Mean of <i>AQI</i>	36.51	36.51	35.7	35.7	35.5	35.5
Mean of <i>VP</i>	—	0.31	—	0.16	—	0.2
Mean of <i>VE</i>	—	0.48	—	0.4	—	0.3
<i>Panel B. By race (Poisson GMM-IV)</i>						
	Black/Hispanic		White			
	(1)	(2)	(3)	(4)		
<i>AQI</i>	0.024 (0.012)	0.086 (0.035)	0.040 (0.0070)	0.13 (0.023)		
<i>AQI</i> × <i>VP</i>		−0.43 (0.20)		−0.56 (0.13)		
Observations	7,740	7,740	15,553	15,553		
Mean of outcome	3.27	3.27	4.17	4.17		
Mean of <i>AQI</i>	37.5	37.5	35.46	35.46		
Mean of <i>VP</i>	—	0.21	—	0.23		
Mean of <i>VE</i>	—	0.36	—	0.37		

*Notes:* The dependent variable is the count of inpatient hospital admissions with influenza as primary diagnosis within a county-year-month. The columns indicate which age (panel A) or race (panel B) subgroups are counted in the dependent variable. We limit analysis to the influenza-intensive months of October through March, and our sample spans 2007–2017 with the exception of October 2008 to March 2009, where vaccine effectiveness data are not available. Vaccine protection (*VP*) is weighted by hospitalization shares across age groups and is measured between 0 (low) and 1 (high). We only use the vaccine take-up rates and raw vaccine effectiveness for the age groups indicated in each column in panel A. For the results by racial group in panel B, we use our *VP* scaled by the ratio of race specific to overall vaccine take-up by season. The results are from Poisson GMM-IV estimations with county-by-season-by-year fixed effects and year-by-month dummies as well as weather controls. Weather controls consist of five bins of temperature quintiles, five bins of specific humidity quintiles, and linear terms for precipitation and wind speed. All weather variables are based on county-year-month averages. The air quality index (*AQI*) is lagged one month, and a higher *AQI* means worse air quality. The results use our instruments based on wind direction instead of the *AQI* to generate moment conditions, and even-numbered columns additionally use our *VE* instrument instead of *VP* to form moment conditions. The number of included observations can vary across different outcomes due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level.

the more vulnerable groups.<sup>32</sup> Patterns across the youngest and oldest groups are similar to each other and consistent with our main results. The interaction with vaccine protection for the middle age group, however, is imprecise and positive. A positive point estimate on the interaction term implies that vaccines do not help reduce influenza hospitalizations due to air pollution but can still reduce influenza hospitalizations not driven by air pollution. The confidence intervals are large, however, and overlap with the confidence intervals of the other age groups, so we draw little inference from this age group estimate.

<sup>32</sup>We define these age splits based on the age splits available in the vaccine effectiveness measures.

TABLE 3—THE EFFECT OF AIR POLLUTION AND VACCINES ON EMERGENCY DEPARTMENT VISITS AND MORTALITY

	ED visits				Mortality			
	Poisson GMM		Poisson GMM-IV		Poisson GMM		Poisson GMM-IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>AQI</i>	0.018 (0.0027)	0.059 (0.010)	0.038 (0.0071)	0.11 (0.019)	0.011 (0.0023)	0.028 (0.0073)	0.0014 (0.0080)	0.053 (0.029)
<i>AQI</i> × <i>VP</i>		-0.22 (0.047)		-0.43 (0.11)		-0.088 (0.036)		-0.30 (0.15)
Observations	10,049	10,049	10,049	10,049	23,126	23,126	23,126	23,126
Mean of outcome	38.4	38.4	38.4	38.4	0.96	0.96	0.96	0.96
Mean of <i>AQI</i>	35.3	35.3	35.3	35.3	37.41	37.41	37.41	37.41
Mean of <i>VP</i>	—	0.21	—	0.21	—	0.2	—	0.2
Mean of <i>VE</i>	—	0.37	—	0.37	—	0.35	—	0.35

*Notes:* The dependent variable is the count of ED visits (in columns 1 to 4) or the count of deaths (in columns 5 to 6), all with influenza as primary diagnosis within a county-year-month. We limit analysis to the influenza-intensive months of October through March, and our sample spans 2007–2017 with the exception of October 2008 to March 2009, where vaccine effectiveness data are not available. Vaccine protection (*VP*) is weighted by hospitalization shares across age groups and is measured between 0 (low) and 1 (high). We only use the vaccine take-up rates and raw vaccine effectiveness for the age groups indicated in each column. For the results by racial groups, we use our *VP* scaled by the ratio of race specific to overall vaccine take-up by season. The results are from a Poisson GMM estimation with county-by-season-by-year fixed effects and year-by-month dummies as well as weather controls. Weather controls consist of five bins of temperature quintiles, five bins of specific humidity quintiles, and linear terms for precipitation and wind speed. All weather variables are based on county-year-month averages. The air quality index (*AQI*) is lagged one month, and a higher *AQI* means worse air quality. The columns indicating “GMM-IV” use our instruments based on wind direction instead of the *AQI* to generate moment conditions, and even-numbered columns additionally use our *VE* instrument instead of *VP* to form moment conditions. The number of included observations can vary across different outcomes due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level.

Estimates are similar across racial and ethnic groups (Blacks/Hispanics and Whites in columns 1 to 4 in panel B), with overlapping confidence intervals.<sup>33</sup> Combining these results with well-established racial and ethnic differences in pollution exposure (Banzhaf, Ma, and Timmins 2019; Colmer et al. 2020; Currie, Voorheis, and Walker 2020) may help explain the higher influenza burdens experienced by those communities (e.g., Quinn et al. 2011). As such, our results suggest that air quality control could be an additional policy lever to help reduce severe influenza cases among these vulnerable groups, particularly within those communities in which vaccine access is limited and reluctance to receive the vaccine is particularly high.<sup>34</sup>

Although we focus primarily on inpatient hospital admissions for influenza, Table 3 shows estimates of the effect of air pollution and vaccines on two alternative outcomes: ED visits and mortality. ED visits may pick up less severe cases of the flu, though visiting the ED can be plagued by selection concerns since they are more likely to serve as a source of primary care for groups with limited access to health care (Finkelstein et al. 2012). Despite the fact that our data on ED visits have slightly different geographical and temporal coverage than the data for inpatient

<sup>33</sup> We adjust vaccine protection by the seasonal ratio of vaccine take-up of the particular ethnic group to overall vaccine take-up, which results in a slightly higher mean of *VP* for Whites, as reported in the bottom of the table.

<sup>34</sup> These benefits are in addition to any improvements in pollution-related health not associated with influenza. See Deryugina et al. (2021) for a discussion of policy targeting regarding polluted areas and vulnerable people.



hospitalizations, the estimates are close to our main results. In columns 5 to 8, we instead look at influenza deaths, which are less frequent than inpatient hospitalizations but also less subject to selection concerns.<sup>35</sup> The estimates for mortality also show a similar pattern to our main results. Together, these suggest that air pollution and the protective role of vaccines each affect a wide range of flu case severity.

In Table 4 we perform three further tests. First, columns 1 and 2 explore robustness to functional form by using a linear mean function, closer to standard OLS, in place of the exponential mean function consistent with a Poisson count model. Columns 1 and 2 in panel A show a linear GMM model without instruments (OLS), for a linear version of columns 1 and 2 in panel A of Table 1. Columns 1 and 2 in panel B show the IV as a linear version of columns 1 and 2 in panel B of Table 1. As in our baseline Poisson GMM model, the IV estimates in panel B are around three times larger than those in panel A. Since the point estimates now reflect level effects, we divide by the mean of the dependent variable to obtain percent effects that are more readily comparable to the estimates from the count model. Doing so, the linear estimate in column 1 in panel B of 0.18 translates to a 3 percent effect, which is very close to the estimate of 2.8 percent using the count model. Vaccine protection is also comparable in magnitude. In online Appendix Table A.8, we show equivalence of our Poisson GMM estimator (without instruments) with a PPML estimator, and we estimate a linear model using the inverse hyperbolic sine (IHS) of hospitalizations as our outcome. The estimates using the IHS are similar to semielasticities (but, unlike the log function, allow for zeros) and can therefore be more directly compared with our baseline Poisson GMM estimates. The effect of 0.02 in column 7 in online Appendix Table A.8 is close to our baseline effect of 0.028 in Table 1. Together, these results suggest that our estimates are largely insensitive to the functional form choice of our dependent variable.

Second, we ask how the effect of air pollution on influenza hospitalization compares to the effect on any respiratory hospitalization (including influenza) in columns 3 and 4. Here we continue to use the linear models from columns 1 and 2, which has the benefit of making direct level effect comparisons, unlike the Poisson model, which provides relative percentage effects. As indicated in Table 4, the mean of hospitalizations with any respiratory hospitalization per county-year-month (141.32) is much higher than for influenza hospitalizations alone (6.04). Panel A shows the effect on all respiratory hospitalizations without instruments, and panel B with instruments.<sup>36</sup> The absolute effect of a 1-unit increase of the AQI on influenza hospitalizations (0.18, column 1 in panel B) is roughly one-third of the size of the effect on all respiratory hospitalizations (0.5, column 3 in panel B).<sup>37</sup> Assuming that outside of influenza season the effect on all respiratory hospitalizations remains the same but the effect on influenza hospitalizations drops to zero, influenza hospitalizations due to air pollution account

<sup>35</sup> Since the data on mortality cover the entire United States, these results also improve the representativeness of our main findings. The estimation sample size reported in the table is only slightly higher than for our main results because the mortality outcome has more zeros, resulting in more observations being dropped by the count model.

<sup>36</sup> A negative but noisy interaction effect on all respiratory hospitalizations in column 4 in panel B is consistent with our vaccine effects, as the majority of hospitalizations for respiratory diseases are unrelated to influenza (e.g., asthma).

<sup>37</sup> This is further corroborated when comparing the effect of air pollution on influenza with all other diagnoses groupings in online Appendix Figure A.6.

TABLE 4—LINEAR SPECIFICATION, ALL RESPIRATORY HOSPITALIZATIONS, AND OSTEOARTHRITIS AS FALSIFICATION TEST

	Influenza hospitalizations		All respiratory hospitalizations		Osteoarthritis	
<i>Panel A. No instruments</i>	Linear GMM		Linear GMM		Poisson GMM	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>AQI</i>	0.063 (0.025)	0.19 (0.073)	0.17 (0.067)	0.014 (0.16)	-0.00054 (0.00027)	0.00019 (0.00084)
<i>AQI</i> × <i>VP</i>		-0.61 (0.28)		0.72 (0.69)		-0.0034 (0.0041)
<i>Panel B. With instruments</i>	Linear GMM-IV		Linear GMM-IV		Poisson GMM-IV	
<i>AQI</i>	0.18 (0.058)	0.51 (0.16)	0.50 (0.29)	0.33 (0.38)	-0.0016 (0.0014)	0.00069 (0.00029)
<i>AQI</i> × <i>VP</i>		-1.88 (0.79)		-1.32 (2.17)		-0.015 (0.015)
Observations	17,668	17,668	24,596	24,596	24,255	24,255
Mean of outcome	6.04	6.04	141.32	141.32	43.51	43.51
Mean of <i>AQI</i>	35.27	35.27	34.52	34.52	34.54	34.54
Mean of <i>VP</i>	—	0.21	—	0.21	—	0.21
Mean of <i>VE</i>	—	0.36	—	0.37	—	0.37

*Notes:* The dependent variable is the count of inpatient hospitalizations with influenza as primary diagnosis in columns 1 and 2, the count of inpatient hospitalizations with any respiratory primary diagnosis in columns 3 and 4, and the count of inpatient hospitalizations with osteoarthritis as primary diagnosis in columns 5 and 6, all at the county-year-month level. We limit analysis to the influenza-intensive months of October through March, and our sample spans 2007–2017 with the exception of October 2008 to March 2009, where vaccine effectiveness data are not available. Vaccine protection (*VP*) is weighted by hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a linear GMM estimation in columns 1 to 4 and from a Poisson GMM estimation in columns 5 and 6, all with county-by-season-by-year fixed effects and year-by-month dummies as well as weather controls. Weather controls consist of five bins of temperature quintiles, five bins of specific humidity quintiles, and linear terms for precipitation and wind speed. All weather variables are based on county-year-month averages. The air quality index (*AQI*) is lagged one month, and a higher *AQI* means worse air quality. Panel B uses our instruments based on wind direction instead of the *AQI* to generate moment conditions, and even-numbered columns additionally use our *VE* instrument instead of *VP* to form moment conditions. The number of included observations can vary across different outcomes due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level.

for roughly 18 percent of all respiratory hospitalizations due to air pollution. This suggests that the increased incidence of influenza accounts for a sizable share of the health harms from air pollution. It also implies that greater vaccine strain matches and increased take-up rates can reduce a sizable share of hospitalizations from air pollution.

Third, as a general specification test for our model, we perform a falsification test by repeating our analysis using an outcome we do not expect to be related to pollution or vaccines. We choose to narrow our focus to osteoarthritis, which is unlikely to be related to short-term variation in pollution. Our Poisson GMM-IV results in column 5 and 6 in panel B of Table 4 indicate precise zero coefficients on the effect of *AQI* and the interaction with vaccine protection, lending support to our model specification. As a more comprehensive test, online Appendix Figures A.6, A.7, and A.8 show our Poisson GMM-IV estimates for all ICD disease groupings separately, as long as there

are sufficient number of cases in those diseases.<sup>38</sup> Compared to all other disease outcomes, we find that influenza hospitalizations are most affected by the AQI in relative terms.<sup>39</sup> The vast majority of disease outcomes are associated with a precise zero effect, similar to osteoarthritis. There are some diagnoses where we find that AQI increases hospitalizations (Figure A.6), such as acute bronchitis, perinatal conditions (often includes respiratory conditions), or diabetes (known to increase influenza risk (Allard et al. 2010)). Adjusting for the family-wise error rate for multiple hypothesis testing renders all results insignificant, except those for our influenza outcome.

As an expansion to our wind instrumental variables, we explore an additional source of variation by using inversions in online Appendix Table A.9. In columns 1 and 2, we use only inversions (without using instruments based on wind direction). The coefficients are similar as in our main results in Table 1, with overlapping confidence intervals. We next use both the inversion and wind-based sets of instruments in columns 3 and 4, again with estimates close to our main results.<sup>40</sup> These patterns lend support to the validity of our model design and demonstrate that our IV estimates are not a unique feature of our measure of wind direction in the first stage.

Finally, Table A.10 in online Appendix A.3 explores further robustness of our main Poisson GMM-IV results to changes in control variables, calculation of AQI, or including off-seasonal cases. In columns 1 and 2, we replace our county-by-season-by-year fixed effects with coarser county-by-influenza-season effects. In columns 3 and 4, we drop all weather controls. In columns 5 and 6, we use the full controls and additionally include lagged employment at the county-year-month level to control for economic activity at our level of analysis. In columns 7 and 8, we do not winsorize the AQI, and in columns 9 and 10, we do not spatially interpolate the AQI. In columns 11 and 12, we additionally include all county-year-month cells with positive influenza hospitalization cases. The estimates remain similar to our main estimates.

## B. Medical Costs

Given the above effects, we calculate the additional hospital costs and charges attributable to pollution-associated influenza to assess the costs generated by air pollution and the role of vaccine protection in mitigating those costs.<sup>41</sup> We use hospital

<sup>38</sup>We focus on disease groupings that have at least half as many occurrences as influenza. This implies a threshold mean of at least 3.02 in our outcome variable, half the mean of our influenza outcome (6.04). Including additional low-occurrence diseases confirms the pattern shown in the figures.

<sup>39</sup>Note that our estimate is in relative terms due to the exponential mean function. It is possible that the estimate in absolute terms is higher for other diseases, such as acute bronchitis, due to higher baseline prevalence.

<sup>40</sup>The test of overidentifying restrictions is rejected at the 5 percent level, both when using inversion instruments alone and when using inversions and wind instruments jointly. The test for overidentifying restrictions is passed only with instruments based on wind direction alone, as in our main results. This together with the lower first-stage *F*-stat for inversion instruments drives using solely wind direction instruments as main results.

<sup>41</sup>We use deflated hospital charges with base year 2018 and combine them with the HCUP (2018b)-provided cost-to-charge ratios to convert them into hospital costs. Hospital charges are around US\$29,000 per patient per influenza-diagnosed inpatient hospitalization, but actual costs to hospitals are lower at around US\$8,000 per patient (online Appendix Table A.2). Further, these estimates ignore indirect costs to patients, such as forgone earnings.

costs as the dependent variable in Table 5, showing only results using instruments, with noninstrumented results presented in online Appendix Table A.11 and results using hospital charges instead of costs in online Appendix Table A.12. Column 1 in panel A indicates that a 1-unit increase in the AQI corresponds to a US\$1,595 increase in hospital costs from hospitalizations with primary-diagnosis influenza. This implies that a one standard deviation decrease in AQI (10.9 points) reduces hospitalization costs by US\$328 million per influenza season across the entire United States. Column 2 in panel A shows the interaction effect with vaccine protection, and columns 1 and 2 in panel B use a Poisson model instead of a linear model. Since the effect is a relative effect in panel B, the estimates are reassuringly close to our estimates in columns 1 and 2 in panel B in Table 1 where we use the count of hospitalizations as outcome.<sup>42</sup>

We can use the results in column 2 in panel A to further illustrate our main results in terms of additional hospital costs. When *VP* is high (maximum is 0.33), an increase in AQI has no noticeable impact on flu-specific hospitalization costs due to the protective nature of the vaccine. In contrast, when *VP* is low, even small changes in the AQI generate large increases in additional influenza-specific hospitalization costs. Going from an AQI of 40 to 50 (both of which are well below US regulatory standards) generates roughly US\$455 million in additional influenza inpatient hospitalization costs at a vaccine protection of 0.086, the minimum in our sample.<sup>43</sup> Conversely, when air quality is high (*AQI* < 20), a drop in *VP* generates little additional pollution-driven influenza hospitalization costs (though influenza cases that are not pollution driven still might be greatly affected). On the other hand, when air quality approaches an AQI of 70 (which is still relatively clean by WHO standards), *VP* is highly impactful. In particular, a drop in vaccine protection from its median (0.21) to the twenty-fifth percentile (0.16), generates around US\$166 million in additional pollution-driven influenza costs when AQI is at the low end of our sample range and around US\$580 million at the high end of the pollution range.<sup>44</sup>

We also decompose the effect on total costs into two of its three components, the effect on length of stay in days (columns 3 and 4 of Table 5) and costs per day (columns 5 and 6).<sup>45</sup> Examining these two outcomes may also help to shed some light on whether pollution is likely to increase the severity or spread of flu. Our results show a positive but statistically insignificant increase in length of stay and a statistically significant increase in costs per day. We interpret these results as supporting the idea that pollution leads to more intense cases of the flu—i.e., it increases severity. We note, however, that we cannot properly disentangle the spread versus severity story, as we only observe the joint outcome of likelihood of hospitalization and not the two

<sup>42</sup>Note that we have an equally high share of zeros regardless of whether we use hospital costs or count of hospitalizations as our outcome measure. The better fit of an exponential (Poisson) model to data with large shares of zeros may explain why the Poisson based estimates in panel B are slightly more precise than those based on the linear model in panel A.

<sup>43</sup>Calculated as  $10 \cdot (3,117 - 0.08 \cdot 8,794) = \text{US}\$24,135$  per county-month, multiplied by 3,142 counties and 6 months.

<sup>44</sup>Calculated as  $20 \cdot (0.05 \cdot 8,794) = 8,794$  and  $70 \cdot (0.05 \cdot 8,794) = 30,779$ , both multiplied by 3,142 counties and 6 months.

<sup>45</sup>The third component is the number of hospitalizations per county-month as in Table 4 in columns 1 and 2 in panel B.

TABLE 5—TOTAL HOSPITALIZATION COSTS, LENGTH OF STAY, AND COSTS PER DAY

	Total costs		Length of stay in days		Costs per day	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Linear GMM-IV</i>						
<i>AQI</i>	1,594.5 (504.7)	3,117.3 (1,415.5)	0.017 (0.017)	0.070 (0.046)	15.3 (7.34)	20.2 (19.2)
<i>AQI</i> × <i>VP</i>		-8,794.1 (7,227.4)		-0.27 (0.26)		-21.1 (110.8)
<i>Panel B. Poisson GMM-IV</i>						
<i>AQI</i>	0.024 (0.0086)	0.097 (0.026)	0.0065 (0.0077)	0.058 (0.021)	0.0078 (0.0065)	0.023 (0.018)
<i>AQI</i> × <i>VP</i>		-0.53 (0.17)		-0.32 (0.13)		-0.10 (0.11)
Observations	17,754	17,754	17,783	17,783	17,754	17,754
Mean of outcome	48,011.49	48,011.49	2.64	2.64	1,238.3	1,238.3
Mean of <i>AQI</i>	35.28	35.28	35.29	35.29	35.28	35.28
Mean of <i>VP</i>	—	0.21	—	0.21	—	0.21
Mean of <i>VE</i>	—	0.36	—	0.36	—	0.36

*Notes:* The dependent variable is hospital costs for inpatient hospitalizations with influenza as primary diagnosis, length of stay in days, or costs per day. We limit analysis to the influenza-intensive months of October through March, and our sample spans 2007–2017 with the exception of October 2008 to March 2009, where vaccine effectiveness data are not available. Vaccine protection (*VP*) is weighted by hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a linear GMM estimation in panel A and from a Poisson GMM estimation in panel B, all with county-by-season-by-year fixed effects and year-by-month dummies as well as weather controls. Weather controls consist of five bins of temperature quintiles, five bins of specific humidity quintiles, and linear terms for precipitation and wind speed. All weather variables are based on county-year-month averages. The air quality index (*AQI*) is lagged one month, and a higher *AQI* means worse air quality. All results use our instruments based on wind direction instead of the *AQI* to generate moment conditions, and even-numbered columns additionally use our *VE* instrument instead of *VP* to form moment conditions. Standard errors in parentheses are clustered at the county level.

separate components (the likelihood of catching the flu or the likelihood of hospitalization conditional on catching the flu).<sup>46</sup>

A back-of-the-envelope calculation based on realizations within our dataset may help place our estimates in a more useful context. Our results suggest that a 10 percent (3.5 *AQI* points) reduction in the *AQI* in a historically bad vaccine effectiveness year (17 percent *VE* and 11 percent *VP*) would avert 12,607 (16.6 percent) hospitalizations across the United States, or US\$109 million in influenza medical costs.<sup>47</sup> In contrast, a 10 percent improvement in either vaccine take-up or vaccine effectiveness from average vaccine take-up or effectiveness in a historically polluted year (38.2 *AQI*) would avert 26,378 (34.6 percent) of pollution-driven influenza hospitalizations, or US\$211 million.

<sup>46</sup> Admitting more marginally sick people, for example, could undermine this exercise.

<sup>47</sup> We calculate averted medical costs by multiplying the number of averted hospitalizations by the average costs per influenza hospitalization (US\$8,000).

### C. Policy Implications

Since the marginal benefit from improving either *VP* or air quality decreases in the level of the other variable, vaccine and air quality policies can serve as substitutes in preventing pollution-induced influenza cases.<sup>48</sup> The optimal mix of those policies will depend on their relative costs on the margin as well as the spillover impacts each may have on harms beyond influenza. It is also worth noting that these policies can operate on different time scales, adding an additional dimension to the trade-offs across each policy. For example, reducing pollution emissions requires investments in capital equipment and takes considerable time, but informational approaches that promote vaccine take-up can bear fruit much more quickly. Regardless of the specifics, recognizing the interaction of these two policies broadens the toolkit to address the harms from either one of them and thus necessarily allows one to obtain a given set of policy objectives at (weakly) lower costs.

The interaction of vaccine take-up and air quality policies also highlights new potential benefits from improved targeting of either policy. Given the seasonality of the flu, air quality policies could be time-varying.<sup>49</sup> Given our finding that influenza hospitalizations account for a significant share of all respiratory hospitalizations from pollution, more stringent air quality policies in influenza months may be particularly impactful.<sup>50</sup>

Programs to promote vaccine take-up can also be targeted toward vulnerable communities based on sociodemographic risk factors. Using time-averaged data on vaccine take-up by state and race, air quality by county, and sociodemographic county characteristics on median household income and race from Chetty et al. (2018) and the US census, we ask what an increase of 10 percent in the overall vaccine take-up average (a 4.6 percentage point increase) achieves when targeted at different types of counties.<sup>51</sup> Specifically, we examine differences between the top and bottom 1 percent of US counties in terms of racial population shares and median income, and we use our estimated coefficients to compare the benefits of a 4.6 percentage point increase in vaccine take-up across these scenarios.<sup>52</sup>

Increasing vaccine take-up in a county with high versus low Black population shares reduces hospitalizations and associated costs by 40.7 percent versus 26.8 percent. The difference is substantial, since pollution exposure is higher and vaccine take-up lower in communities with higher shares of Black residents. Interestingly, we find little difference in results when stratifying by income. Targeting a high-income versus a low-income county reduces pollution-induced

<sup>48</sup> In the event that air quality improvements reduce the spread of the flu, the two policies may serve as complements rather than substitutes under certain conditions. For example, if the reproductive rate of flu is above 1 even with vaccine protection but falls below 1 when combined with air quality policy, then the two policies may be complements.

<sup>49</sup> We thank Douglas Almond for raising this point.

<sup>50</sup> Seasonal air quality policies are not without precedent. For example, the NO<sub>x</sub> cap-and-trade program in the United States only operates between May and September, when ozone tends to be highest (Deschenes, Greenstone, and Shapiro 2017).

<sup>51</sup> We calculate the average vaccine take-up by county using vaccine take-up rates by race and White, Black, Hispanic, and Asian population shares by county. Our calculations account for the joint statistical relationships between vaccine take up, race, pollution exposure, and income.

<sup>52</sup> We use the median county AQI and vaccine take-up rate in the bottom and top 1 percent of counties.



influenza hospitalizations and associated costs by 36.4 percent versus 38.0 percent. Correlations in our data suggest that these effects are similar due to two offsetting forces. High-income counties tend to include major cities, which are more polluted than more rural areas (increasing the benefits of vaccine-take up) but also have higher baseline vaccine take-up rates (decreasing the benefits). To some degree, this supports the conjecture that air quality controls and vaccine take-up can serve as substitutes.

#### IV. Conclusion

Using a rich, longitudinal dataset, we provide evidence that air pollution increases seasonal influenza hospitalization rates and that improved vaccine protection, either through high vaccine effectiveness or vaccine take-up, greatly diminishes this relationship and reduces the social and medical costs of poor air quality. Our empirical strategy, based on instrumental variables using wind direction and the stochastic nature of vaccine effectiveness across influenza seasons, limits risks of confounding. Our results are robust to numerous assumptions about functional form, omitted variables, alternative outcomes, and falsification tests.

That policies to combat air quality can protect citizens from the most serious threats of influenza is a new insight that offers an additional tool in the global battle against the flu. At the same time, it appears that increased flu vaccination rates and improvements in flu vaccine strain matches can avert some of the harms from pollution. As such, the returns to policies designed to address pollution and infection externalities are inextricably connected, such that approaching either in isolation will be suboptimal from a social welfare perspective. Thus, optimal policy strategies can help decrease medical spending, avoid lost productivity, and reduce loss of life. These returns may be particularly high in dense urban centers around the world and in developing countries in particular, where population density and high levels of pollution (de Lataillade, Auvergne, and Delannoy 2009) increase the intensity of these interactions.

Our insights regarding compounding risks from pollution and flu may extend to other viral respiratory illnesses with similar etiological pathways. For example, research on the COVID-19 pandemic suggests significant effects of pollution on COVID-19 cases and deaths (Wu et al. 2020; Ispording and Pestel 2021).<sup>53</sup> While vaccine developments aid against such health threats, new strains and viruses may emerge that diminish such protection. Our results suggest an additional possible policy direction, whereby environmental controls serve as an investment to optimally manage the harms from new viral threats if effective vaccines are not available while also providing additional protection against more established respiratory infections that may drain the health-care system during times of crises.

<sup>53</sup> See, e.g., Cui et al. (2003) for evidence on the earlier SARS-CoV.

## REFERENCES

- Allard, Robert, Pascale Leclerc, Claude Tremblay, and Terry-Nan Tannenbaum.** 2010. "Diabetes and the Severity of Pandemic Influenza A (H1N1) Infection." *Diabetes Care* 33 (7): 1491–93.
- Arceo, Eva, Rema Hanna, and Paulina Oliva.** 2016. "Does the Effect of Pollution on Infant Mortality Differ between Developing and Developed Countries? Evidence from Mexico City." *Economic Journal* 126 (591): 257–80.
- Banzhaf, Spencer, Lala Ma, and Christopher Timmins.** 2019. "Environmental Justice: The Economics of Race, Place, and Pollution." *Journal of Economic Perspectives* 33 (1): 185–208.
- Belongia, Edward A., Burney A. Kieke, James G. Donahue, Laura A. Coleman, Stephanie A. Irving, Jennifer K. Meece, Mary Vandermause, Stephen Lindstrom, Paul Gargiullo, and David K. Shay.** 2011. "Influenza Vaccine Effectiveness in Wisconsin during the 2007–08 Season: Comparison of Interim and Final Results." *Vaccine* 29 (38): 6558–63.
- Boulier, Bryan L., Tejwant S. Datta, and Robert S. Goldfarb.** 2007. "Vaccination Externalities." *BE Journal of Economic Analysis and Policy* 7 (1): 1–27.
- Brauer, Michael, Gerard Hoek, Patricia Van Vliet, Kees Meliefste, Paul H. Fischer, Alet Wijga, Laurens P. Koopman et al.** 2002. "Air Pollution from Traffic and the Development of Respiratory Infections and Asthmatic and Allergic Symptoms in Children." *American Journal of Respiratory and Critical Care Medicine* 166 (8): 1092–98.
- Bureau of Labor Statistics.** 2021. "Local Area Unemployment Statistics (LAUS)." United States Bureau of Labor Statistics. <https://download.bls.gov/pub/time.series/la/la.data.64.County> (accessed February 9, 2021).
- Casanova, Lisa M., Soyoung Jeon, William A. Rutala, David J. Weber, and Mark D. Sobsey.** 2010. "Effects of Air Temperature and Relative Humidity on Coronavirus Survival on Surfaces." *Applied and Environmental Microbiology* 76 (9): 2712–17.
- CDC (Centers for Disease Control and Prevention).** 2008. "Behavioral Risk Factor Surveillance System (BRFSS) Survey Data 2007–08." Centers for Disease Control and Prevention. [https://www.cdc.gov/brfss/annual\\_data/annual\\_2008.htm](https://www.cdc.gov/brfss/annual_data/annual_2008.htm) (accessed March 4, 2020).
- CDC.** 2009. "Estimated Proportion of Persons Aged 6 Months Who Received an Influenza Vaccination during the 2008–09 Influenza Season (September 2008 through February 2009), by Age Group, High-Risk Status, and Race/Ethnicity." Centers for Disease Control and Prevention. [www.cdc.gov/vaccines/imz-managers/coverage/nhis/downloads/2009-nhis-tables.xls](http://www.cdc.gov/vaccines/imz-managers/coverage/nhis/downloads/2009-nhis-tables.xls) (accessed March 4, 2020).
- CDC.** 2015. "Cumulative Monthly Influenza Vaccination Coverage Estimates for Persons 6 Months and Older by State, HHS Region, and the United States, National Immunization Survey-Flu (NIS-Flu) and Behavioral Risk Factor Surveillance System (BRFSS), 2009–10 Influenza Season." Centers for Disease Control and Prevention. <https://data.cdc.gov/Flu-Vaccinations/Influenza-Vaccination-Coverage-for-All-Ages-6-Mont/vh55-3he6> (accessed March 4, 2020).
- CDC.** 2019. *Seasonal Flu Vaccine Effectiveness Studies*. Centers for Disease Control and Prevention.
- CDC.** 2020. "Influenza Vaccination Coverage Estimates for Persons 6 Months and Older by State, HHS region, and the United States, National Immunization Survey-Flu (NIS-Flu) and Behavioral Risk Factor Surveillance System (BRFSS), 2010–11 through 2018–19 Influenza Seasons." Centers for Disease Control and Prevention. <https://data.cdc.gov/Flu-Vaccinations/Influenza-Vaccination-Coverage-for-All-Ages-6-Mont/vh55-3he6> (accessed March 4, 2020).
- Chen, Pei-Shih, Feng Ta Tsai, Chien Kun Lin, Chun-Yuh Yang, Chang-Chuan Chan, Chea-Yuan Young, and Chien-Hung Lee.** 2010. "Ambient Influenza and Avian Influenza Virus during Dust Storm Days and Background Days." *Environmental Health Perspectives* 118 (9): 1211–16.
- Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter.** 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." NBER Working Paper No. 25147.
- Cienciewicki, Jonathan, and Ilona Jaspers.** 2007. "Air Pollution and Respiratory Viral Infection." *Inhalation Toxicology* 19 (14): 1135–46.
- Clay, Karen, Joshua Lewis, and Edson Severnini.** 2018. "Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic." *Journal of Economic History* 78 (4): 1179–1209.
- Cohen, Jon.** 2017. "Why Flu Vaccines So Often Fail." *Science*, September 20. <https://www.science.org/content/article/why-flu-vaccines-so-often-fail>.
- Cohen, Aaron J., Michael Brauer, Richard Burnett, H. Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan et al.** 2017. "Estimates and 25-Year Trends of the Global Burden of Disease Attributable to Ambient Air Pollution: An Analysis of Data from the Global Burden of Diseases Study 2015." *The Lancet* 389 (10082): 1907–18.

- Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis.** 2020. "Disparities in PM2.5 Air Pollution in the United States." *Science* 369 (6503): 575–78.
- Correia, Sergio, Paulo Guimarães, and Thomas Zylkin.** 2019. "PPMLHDFE: Fast Poisson Estimation with High-Dimensional Fixed Effects." *arXiv*: 1903.01690.
- Cui, Yan, Zuo-Feng Zhang, John Froines, Jinkou Zhao, Hua Wang, Shun-Zhang Yu, and Roger Detels.** 2003. "Air Pollution and Case Fatality of SARS in the People's Republic of China: An Ecologic Study." *Environmental Health* 2 (1): 1–5.
- Currie, Janet, John Voorheis, and Reed Walker.** 2020. "What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality." NBER Working Paper No. 26659.
- de Lataillade, Camille, Stéphane Auvergne, and Isabelle Delannoy.** 2009. "2005 and 2006 Seasonal Influenza Vaccination Coverage Rates in 10 Countries in Africa, Asia Pacific, Europe, Latin America and the Middle East." *Journal of Public Health Policy* 30 (1): 83–101.
- Deryugina, Tatyana, Nolan Miller, David Molitor, and Julian Reif.** 2021. "Geographic and Socioeconomic Heterogeneity in the Benefits of Reducing Air Pollution in the United States." *Environmental and Energy Policy and the Economy* 2 (1): 157–89.
- Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif.** 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review* 109 (12): 4178–4219.
- Deschenes, Olivier, Michael Greenstone, and Joseph S. Shapiro.** 2017. "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." *American Economic Review* 107 (10): 2958–89.
- Diamond, Gill, Diana Legarda, and Lisa K. Ryan.** 2000. "The Innate Immune Response of the Respiratory Epithelium." *Immunological Reviews* 173: 27–38.
- Din, Alexander, and Ron Wilson.** 2020. "Crosswalking ZIP Codes to Census Geographies." *Cityscape* 22 (1): 293–314.
- EPA (Environmental Protection Agency).** 2018. *Technical Assistance Document for the Reporting of Daily Air Quality*. Washington, DC: United States Environmental Protection Agency.
- EPA.** 2020. "Air Quality System Data Mart." United States Environmental Protection Agency. [https://aqs.epa.gov/aqsweb/airdata/download\\_files.html](https://aqs.epa.gov/aqsweb/airdata/download_files.html) (accessed March 4, 2020).
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen et al.** 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *Quarterly Journal of Economics* 127 (3): 1057–1106.
- Flannery, Brendan, Jessie R. Chung, Arnold S. Monto, Emily T. Martin, Edward A. Belongia, Huong Q. McLean, Manjusha Gaglani et al.** 2019. "Influenza Vaccine Effectiveness in the United States during the 2016–2017 Season." *Clinical Infectious Diseases* 68 (11): 1798–1806.
- Flannery, Brendan, Rebecca J. Garten Kondor, Jessie R. Chung, Manjusha Gaglani, Michael Reis, Richard K. Zimmerman, Mary Patricia Nowalk et al.** 2020. "Spread of Antigenically Drifted Influenza A(H3N2) Viruses and Vaccine Effectiveness in the United States during the 2018–2019 Season." *Journal of Infectious Diseases* 221 (1): 8–15.
- Francis, Thomas.** 1960. "On the Doctrine of Original Antigenic Sin." *Proceedings of the American Philosophical Society* 104 (6): 572–78.
- Gaglani, Manjusha, Jessica Pruszynski, Kempapura Murthy, Lydia Clipper, Anne Robertson, Michael Reis, Jessie R. Chung et al.** 2016. "Influenza Vaccine Effectiveness against 2009 Pandemic Influenza A(H1N1) Virus Differed by Vaccine Type during 2013–2014 in the United States." *Journal of Infectious Diseases* 213 (10): 1546–56.
- GMAO (Global Modeling and Assimilation Office).** 2015. "MERRA-2 inst6\_3d\_ana\_Np: 3d,6-Hourly, Instantaneous, Pressure-Level, Analysis, Analyzed Meteorological Fields V5.12.4." GMAO. [https://disc.gsfc.nasa.gov/datasets/M2I6NPANA\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I6NPANA_5.12.4/summary) (accessed March 4, 2020).
- Gourieroux, Christian, Alain Monfort, and Alain Trognon.** 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica* 52 (3): 701–20.
- Graff Zivin, Joshua, Matthew Neidell, Nicholas J. Sanders, and Gregor Singer.** 2023. "Replication data for: When Externalities Collide: Influenza and Pollution." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.38886/E165401V1>.
- Greenburg, Leonard, Franklyn Field, Carl L. Erhardt, Marvin Glasser, and Joseph I. Reed.** 1967. "Air Pollution, Influenza, and Mortality in New York City: January–February 1963." *Archives of Environmental Health: An International Journal* 15 (4): 430–38.

- Griffin, Marie R., Arnold S. Monto, Edward A. Belongia, John J. Treanor, Qingxia Chen, Jufu Chen, H. Keipp Talbot et al. 2011. "Effectiveness of Non-adjuvanted Pandemic Influenza A Vaccines for Preventing Pandemic Influenza Acute Respiratory Illness Visits in 4 US Communities." *PLoS ONE* 6 (8): 1–7.
- Hahon, Nicholas, James A. Booth, Francis Green, and Trent R. Lewis. 1985. "Influenza Virus Infection in Mice after Exposure to Coal Dust and Diesel Engine Emissions." *Environmental Research* 37 (1): 44–60.
- Han, Lefei, Jinjun Ran, Yim-Wah Mak, Lorna Kwai-Ping Suen, Paul H. Lee, Joseph Sriyal Malik Peiris, and Lin Yang. 2019. "Smoking and Influenza-Associated Morbidity and Mortality: A Systematic Review and Meta-analysis." *Epidemiology* 30 (3): 405–17.
- Harper, G.J. 1961. "Airborne Micro-organisms: Survival Tests with Four Viruses." *Epidemiology and Infection* 59 (4): 479–86.
- HCUP (Healthcare Cost and Utilization Project). 2018a. "HCUP State Emergency Department Databases (SEDD)." HCUP. [https://www.hcup-us.ahrq.gov/tech\\_assist/centdist.jsp](https://www.hcup-us.ahrq.gov/tech_assist/centdist.jsp) (accessed September 3, 2020).
- HCUP. 2018b. "HCUP State Inpatient Databases (SID)." HCUP. [https://www.hcup-us.ahrq.gov/tech\\_assist/centdist.jsp](https://www.hcup-us.ahrq.gov/tech_assist/centdist.jsp) (accessed September 3, 2020).
- Hsiang, Solomon, Paulina Oliva, and Reed Walker. 2019. "The Distribution of Environmental Damages." *Review of Environmental Economics and Policy* 13 (1): 83–103.
- Ijaz, M. Khalid, Syed A. Sattar, C. Margaret Johnson-Lussenburg, and V.S. Springthorpe. 1985. "Comparison of the Airborne Survival of Calf Rotavirus and Poliovirus Type 1 (Sabin) Aerosolized as a Mixture." *Applied and Environmental Microbiology* 49 (2): 289–93.
- Isphording, Ingo E., and Nico Pestel. 2021. "Pandemic Meets Pollution: Poor Air Quality Increases Deaths by COVID-19." *Journal of Environmental Economics and Management* 108: 1–25.
- Iuliano, A. Danielle, Katherine M. Roguski, Howard H. Chang, David J. Muscatello, Rakhee Palekar, Stefano Tempia, Cheryl Cohen et al. 2018. "Estimates of Global Seasonal Influenza-Associated Respiratory Mortality: A Modelling Study." *Lancet* 391 (10127): 1285–1300.
- Jackson, Michael L., Jessie R. Chung, Lisa A. Jackson, C. Hallie Phillips, Joyce Benoit, Arnold S. Monto, Emily T. Martin et al. 2017. "Influenza Vaccine Effectiveness in the United States during the 2015–2016 Season." *New England Journal of Medicine* 377 (6): 534–43.
- Jaspers, Ilona, Jonathan M. Ciencewicky, Wenli Zhang, Luisa E. Brighton, Johnny L. Carson, Melinda A. Beck, and Michael C. Madden. 2005. "Diesel Exhaust Enhances Influenza Virus Infections in Respiratory Epithelial Cells." *Toxicological Sciences* 85 (2): 990–1002.
- Khare, P., and L.C. Marr. 2015. "Simulation of Vertical Concentration Gradient of Influenza Viruses in Dust Resuspended by Walking." *Indoor Air* 25 (4): 428–40.
- Kremer, Michael, and Heidi Williams. 2010. "Incentivizing Innovation: Adding to the Tool Kit." *Innovation Policy and the Economy* 10 (1): 1–17.
- Lambert, Linda C., and Anthony S. Fauci. 2010. "Influenza Vaccines for the Future." *New England Journal of Medicine* 363 (21): 2036–44.
- Lee, Greg I., Jordy Saravia, Dahui You, Bishwas Shrestha, Sridhar Jaligama, Valerie Y. Hebert, Tammy R. Dugas, and Stephania A. Cormier. 2014. "Exposure to Combustion Generated Environmentally Persistent Free Radicals Enhances Severity of Influenza Virus Infection." *Particle and Fibre Toxicology* 11 (1): 57.
- Liang, Yijia, Liqun Fang, Hui Pan, Kezhong Zhang, Haidong Kan, Jeffrey R. Brook, and Qinghua Sun. 2014. "PM<sub>2.5</sub> in Beijing—Temporal Pattern and Its Association with Influenza." *Environmental Health* 13 (1): 102.
- Lou, Cairong, Hongyu Liu, Yufeng Li, Yan Peng, Juan Wang, and Lingjun Dai. 2017. "Relationships of Relative Humidity with PM<sub>2.5</sub> and PM<sub>10</sub> in the Yangtze River Delta, China." *Environmental Monitoring and Assessment* 189 (11): 582.
- Lowen, Anice C., Samira Mubareka, John Steel, and Peter Palese. 2007. "Influenza Virus Transmission is Dependent on Relative Humidity and Temperature." *PLoS Pathogens* 3 (10): 1470–76.
- Lu, Peng-Jun, James A. Singleton, Gary L. Euler, Walter W. Williams, and Carolyn B. Bridges. 2013. "Seasonal Influenza Vaccination Coverage among Adult Populations in the United States, 2005–2011." *American Journal of Epidemiology* 178 (9): 1478–87.
- McLean, Huang Q., Mark G. Thompson, Maria E. Sundaram, Burney A. Kieke, Manjusha Gaglani, Kempapura Murthy, Pedro A. Piedra et al. 2015. "Influenza Vaccine Effectiveness in the United States during 2012–2013: Variable Protection by Age and Virus Type." *Journal of Infectious Diseases* 211 (10): 1529–40.

- Mocko, David, and NASA/GSFC/HSL.** 2012. "NLDAS Primary Forcing Data L4 Monthly 0.125 x 0.125 Degree V002." Goddard Earth Sciences Data and Information Services Center. [https://disc.gsfc.nasa.gov/datasets/NLDAS\\_FORA0125\\_M\\_002/summary](https://disc.gsfc.nasa.gov/datasets/NLDAS_FORA0125_M_002/summary) (accessed March 4, 2020).
- Mullins, Jamie T., and Corey White.** 2020. "Can Access to Health Care Mitigate the Effects of Temperature on Mortality?" *Journal of Public Economics* 191: 104259.
- National Center for Health Statistics (NCHS).** 2019. "Vital Statistics Data: 1989–2018 Detailed Mortality." NCHS. <https://www.cdc.gov/nchs/nvss/nvss-restricted-data.htm> (accessed September 3rd, 2020).
- Ohmit, Suzanne E., Mark G. Thompson, Joshua G. Petrie, Swathi N. Thaker, Michael L. Jackson, Edward A. Belongia, Richard K. Zimmerman et al.** 2014. "Influenza Vaccine Effectiveness in the 2011–2012 Season: Protection against Each Circulating Virus and the Effect of Prior Vaccination on Estimates." *Clinical Infectious Diseases* 58 (3): 319–27.
- Putri, Wayan C.W.S., David J. Muscatello, Melissa S. Stockwell, and Anthony T. Newall.** 2018. "Economic Burden of Seasonal Influenza in the United States." *Vaccine* 36 (27): 3960–66.
- Quinn, Sandra Crouse, Supriya Kumar, Vicki S. Freimuth, Donald Musa, Nestor Casteneda-Angarita, and Kelley Kidwell.** 2011. "Racial Disparities in Exposure, Susceptibility, and Access to Health Care in the US H1N1 Influenza Pandemic." *American Journal of Public Health* 101 (2): 285–93.
- Rivas-Santiago, César E., Srijata Sarkar, Pasquale Cantarella, Álvaro Osornio-Vargas, Raúl Quintana-Belmares, Qingyu Meng, Thomas J. Kirm et al.** 2015. "Air Pollution Particulate Matter Alters Antimycobacterial Respiratory Epithelium Innate Immunity." *Infection and Immunity* 83 (6): 2507–17.
- Rolfes, Melissa A., Brendan Flannery, Jessie R. Chung, Alissa O'Halloran, Shikha Garg, Edward A. Belongia, Manjusha Gaglani et al.** 2019. "Effects of Influenza Vaccination in the United States during the 2017–2018 Influenza Season." *Clinical Infectious Diseases* 69 (11): 1845–53.
- Schiller, Jeannine S., and G.L. Euler.** 2009. *Vaccination Coverage Estimates from the National Health Interview Survey: United States, 2008*. Atlanta, GA: Centers for Disease Control and Prevention.
- Shaman, Jeffrey, and Melvin Kohn.** 2009. "Absolute Humidity Modulates Influenza Survival, Transmission, and Seasonality." *PNAS* 106 (9): 3243–48.
- Shaman, Jeffrey, Virginia E. Pitzer, Cécile Viboud, Bryan T. Grenfell, and Marc Lipsitch.** 2010. "Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States." *PLoS Biology* 8 (2): 1–13.
- Silva, J.M.C. Santos, and Silvana Tenreiro.** 2006. "The Log of Gravity." *Review of Economics and Statistics* 88 (4): 641–58.
- Silva, J.M.C. Santos, and Silvana Tenreiro.** 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator." *Economics Letters* 112 (2): 220–22.
- Stoecker, Charles, Nicholas J. Sanders, and Alan Barreca.** 2016. "Success Is Something to Sneeze At: Influenza Mortality in Cities That Participate in the Super Bowl." *American Journal of Health Economics* 2 (1): 125–43.
- Tellier, Raymond.** 2009. "Aerosol Transmission of Influenza A Virus: A Review of Mew Studies." *Journal of the Royal Society Interface* 6: S783–S790.
- Treanor, John J., H. Keipp Talbot, Suzanne E. Ohmit, Laura A. Coleman, Mark G. Thompson, Po-Yung Cheng, Joshua G. Petrie et al.** 2012. "Effectiveness of Seasonal Influenza Vaccines in the United States during a Season with Circulation of All Three Vaccine Strains." *Clinical Infectious Diseases* 55 (7): 951–59.
- Tschofen, Peter, Inês L. Azevedo, and Nicholas Z. Muller.** 2019. "Fine Particulate Matter Damages and Value Added in the US Economy." *PNAS* 116 (40): 19857–62.
- United States Census Bureau.** 2018. *2018 Population Estimates FIPS Codes*. Washington, DC: United States Census Bureau.
- United States Census Bureau.** 2020. "2010 US Census." US Census Bureau. <https://www2.census.gov/programs-surveys/pepost/datasets/> (accessed March 4, 2020).
- Ward, Courtney J.** 2014. "Influenza Vaccination Campaigns: Is an Ounce of Prevention Worth a Pound of Cure?" *American Economic Journal: Applied Economics* 6 (1): 38–72.
- White, Corey.** 2021. "Measuring Social and Externality Benefits of Influenza Vaccination." *Journal of Human Resources* 56 (3): 749–85.
- Wolkoff, Peder.** 2018. "Indoor Air Humidity, Air Quality, and Health – An Overview." *International Journal of Hygiene and Environmental Health* 221 (3): 376–90.
- Wong, Chit Ming, Lin Yang, Thuan Quoc Thach, Patsy Yuen Kwan Chau, King Pan Chan, G. Neil Thomas, Tai Hing Lam et al.** 2009. "Modification by Influenza on Health Effects of Air Pollution in Hong Kong." *Environmental Health Perspectives* 117 (2): 248–53.



- Wu, Xiao, Rachel C. Nethery, Benjamin M. Sabath, Danielle Braun, and Francesca Dominici.** 2020. "Exposure to Air Pollution and COVID-19 Mortality in the United States." *medRxiv*.
- Xia, Youlong, Kenneth Mitchell, Michael Ek, Justin Sheffield, Brian Cosgrove, Eric Wood, Lifeng Luo et al.** 2012. "Continental-Scale Water and Energy Flux Analysis and Validation for the North American Land Data Assimilation System Project Phase 2 (NLDAS-2): 1. Intercomparison and Application of Model Products." *Journal of Geophysical Research: Atmospheres* 117 (D3).
- Zhang, Jun, and Kai F. Yu.** 1998. "What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes." *JAMA* 280 (19): 1690–91.
- Zimmerman, Richard K., Mary Patricia Nowalk, Jessie Chung, Michael L. Jackson, Lisa A. Jackson, Joshua G. Petrie, Arnold S. Monto et al.** 2016. "2014–2015 Influenza Vaccine Effectiveness in the United States by Vaccine Type." *Clinical Infectious Diseases* 63 (12): 1564–73.