NBER WORKING PAPER SERIES

WHEN EXTERNALITIES COLLIDE: INFLUENZA AND POLLUTION

Joshua S. Graff Zivin
Matthew J. Neidell
Nicholas J. Sanders
Gregor Singer

Working Paper 27982
http://www.nber.org/papers/w27982

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020

We thank Max Auffhammer, Luisa Osang, Jeffrey Shaman, and participants at the University of California Environmental Economics seminar for helpful discussions. All errors are our own. The IRB for access to the HCUP data through the National Bureau of Economic Research (NBER) was approved by the NBER. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w27982.ack

When Externalities Collide: Influenza and Pollution
Joshua S. Graff Zivin, Matthew J. Neidell, Nicholas J. Sanders, and Gregor Singer
NBER Working Paper No. 27982
October 2020
JEL No. H23,I12,Q53

ABSTRACT

Influenza and air pollution are significant public health risks with large economic consequences shared across the globe. The common etiological pathways through which they harm health present an interesting case of compounding risk via interacting externalities. Using regional and temporal variation in pollution and disease transmission, we find exposure to more air pollution significantly increases influenza hospitalizations. By exploiting the random deviations in influenza vaccine effectiveness over time, we show high influenza vaccine effectiveness neutralizes this relationship. This suggests seemingly disparate policy actions of pollution control and expanded vaccination provide greater returns than found when studied in isolation.

Joshua S. Graff Zivin
University of California, San Diego
9500 Gilman Drive, MC 0519
La Jolla, CA 92093-0519
and NBER
jgraffzivin@ucsd.edu

Nicholas J. Sanders
Cornell University
Kennedy Hall, Room 422
Ithaca, NY 14853
and NBER
njsanders@cornell.edu

Matthew J. Neidell
Department of Health Policy and Management
Columbia University
722 W 168th Street, 6th Floor
New York, NY 10032
and NBER
mn2191@columbia.edu

Gregor Singer
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0519
gsinger@ucsd.edu

Influenza (flu) and air pollution are both significant public health risks that impact nations around the world. The flu causes an estimated 3-5 million severe cases per year and nearly half a million deaths (Lambert & Fauci 2010, Iuliano et al. 2018), and pollution causes 4.5 million annual deaths (Cohen et al. 2017), with combined annual economic costs estimated to exceed $US 800 billion in the U.S. alone (Putri et al. 2018, Tschofen, Azevedo & Muller 2019). While public health policies to address these issues are often considered in isolation, both share etiological pathways through which they harm human health. We show these externalities interact to influence population health risks, and argue that the stochastic nature of influenza severity and pollution shocks can make public health efforts to combat each complements to one another.

Interactions between the flu and pollution also present an 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 externalities have significant interactive effects: the flu vaccine can protect against 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 demonstrate this interaction in two steps. First, we use hospital data from 2007-2017 across 21 U.S. states to expand upon the cross-sectional epidemiological literature[1] to establish a causal relationship between air pollution and flu cases.[2] We use patient-level data on inpatient hospitalizations, which allows us to focus on cases limited to influenza. We estimate models with spatial and temporal fixed effects to control for numerous unobservable factors. Plausibly exogenous variation in pollution levels within county by quarter-year periods, controlling for seasonality and weather, provides identification. We find pollution significantly increases flu inpatient hospitalizations; a one-standard-deviation increase in the monthly Air Quality Index (10.9-unit increase in our data) amounts to approximately 10.2% additional flu-related inpatient hospitalizations in the U.S. during influenza season.

Second, we exploit data on influenza vaccine effectiveness to explore how the vaccine moderates the estimated relationship above. 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.[3] This makes the random draw of the viral match orthogonal to unobserved determinants of health and air pollution levels, 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 coefficient on pollution, then it

[1] 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 & Severnini (2018). In a study of the Spanish flu in 1918, Clay, Lewis & Severnini (2018) show cities with higher coal-fired power generating capacity saw higher mortality rates, potentially through exposure to higher air pollution.

[2] Air pollution could affect influenza hospitalizations via both susceptibility and exposure. 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 & Ryan 2000, Jaspers et al. 2005, Ciencewicki & Jaspers 2007, Rivas-Santiago et al. 2015). Existing medical research finds exposing *in vitro* respiratory epithelial cells to air pollution increases susceptibility and penetration of influenza (Jaspers et al. 2005), and experimental exposure of mice to air pollution before influenza infections increases morbidity and mortality (Hahon et al. 1985, Lee et al. 2014). Like humidity and temperature (Lowen et al. 2007, Shaman & Kohn 2009, Shaman et al. 2010, Ijaz et al. 1985, Casanova et al. 2010), air pollution particles could also impact the airborne survival of viruses outside the body (Ijaz et al. 1985, Tellier 2009, Chen et al. 2010, Khare & Marr 2015, Lou et al. 2017, Wolkoff 2018) and thus increase the probability of disease transmission.

[3] Other papers using similar variation include Ward (2014) and White (2019).

must be the case that pollution causes influenza hospitalizations.

When we include an interaction between air pollution and vaccine effectiveness, we find the flu vaccine offers significant protection from influenza-related costs of pollution. A vaccine effectiveness of 44%, close to the highest value in our sample, fully neutralizes the relationship between pollution and flu hospitalizations. While this measure of vaccine effectiveness is plausibly exogenous, the nature of its construction allows us to view a percentage change in effectiveness as an equivalent change in the more policy relevant lever of vaccine take-up. Our estimate implies that, at the average vaccine effectiveness of 37%, a 19% increase in vaccine take-up would have the equivalent neutralizing effect.

We also explore results by race and ethnicity. Both of our main findings – that air pollution affects flu hospitalizations and vaccine effectiveness moderates this relationship – are robust across these dimensions. Combined with evidence of significant differences in flu incidence and severity by race (e.g. Quinn et al. 2011), our results highlight that the well-established differences in ambient pollution concentrations across racial and ethnic groups (e.g. Banzhaf, Ma & Timmins 2019, Colmer et al. 2020, Currie, Voorheis & Walker 2020) may serve as an important mechanism driving disparities in influenza outcomes. Moreover, since flu vaccines protect against pollution-induced harms, the private and external benefits from vaccines is considerably higher in communities of color disproportionately exposed to poor air quality.

An important feature of our context is that the spread of influenza and pollution are externalities, in which risks to human harm are stochastic. As externalities, they justify government intervention in the form of policies to increase vaccine take-up and to improve air quality.[4] Insofar as pollution and flu risks have independent variation, policies to address them will be complementary; the variability in pollution levels and vaccine effectiveness that enables our empirical identification ensures that this is true. As such, public health policies, both medical and environmental, can play an important role in hedging against these compounding health risks and their associated economic costs. A back of the envelope calculation suggests a 10% (3.5 AQI points) reduction in the AQI in an historically ineffective vaccine year (17% effectiveness) with an average vaccine take-up would avert 8.2% of all influenza-associated hospitalizations across the U.S. Meanwhile a 10% improvement in vaccine take-up at the average vaccine effectiveness (or, equivalently, a 10% improvement in vaccine effectiveness at the average vaccine take-up) in a historically polluted year (38.2 AQI) would avert 13.2% of influenza hospitalizations.

I. Data

Hospitalizations: We use patient-level data on inpatient hospitalizations from the Health Care and Utilization Project (HCUP 2018*b*). We focus on severe cases specifically limited to influenza using patient level information on diagnosed diseases per International Classification of Diseases (ICD) codes.[5] We use data from 2007 to 2017, for which we also have detailed vaccine effectiveness data available. This covers 21 U.S. states, with an average of 5.5 years per state (see Table A.1 in Appendix A.1 for details on data availability by state and year).

Our principal outcome is the count of inpatient admissions per county-year-month where the ICD code indi-

[4] 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 & Williams 2010).

[5] We exclude patients whose zip code is from a different state than the hospital in which they are treated.

cates influenza.[6] Given the presence of primary and secondary diagnosis codes, we explore three possible ways for classifying flu admissions: (i) limit to admissions where the only diagnosis is influenza (most restrictive); (ii) limit to where any diagnosis is influenza (least restrictive); and (iii) limit to admissions where the primary diagnosis is influenza. The third option reflects a middle ground and is our baseline outcome, and we present results using the other outcomes as well.

We focus on the influenza season (October to March), but also explore results extending the season in Appendix A.2. Figure 1a shows the seasonality of inpatient hospitalizations in our data, which matches closely with general influenza-like illnesses reported by the Centers for Disease Control and Prevention (CDC) (see Table A.1 in 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.[7] During the influenza season from October to March, 54% of county-year-months have no influenza-related hospital admissions in the HCUP data. Our results are robust to inclusion or exclusion of zero valued county-year-months.

As 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.[8] In Appendix A.2, we use outpatient data from emergency departments (HCUP 2018*a*) instead of the inpatient data, with the same strategy of counting patients with a primary diagnosis of influenza as above.

Air quality: We combine hospital admission data with air pollution readings of local ground monitors across the U.S. As our measure of pollution, we use the U.S. Environmental Protection Agency's (EPA 2020) Air Quality Index (AQI) at the daily county level, which we aggregate to county-by-year-by-month to match hospitalization outcomes.[9] We focus on the AQI because it is a summary measure of overall air quality, based on the primary criteria pollutants specified in the Clean Air Act.[10]

Weather controls: We use monthly weather averages from Xia et al. (2012), Mocko & NASA/GSFC/HSL (2012), including temperature, specific humidity, wind speed, and precipitation at the 0.125 by 0.125 degree level, and aggregate up to the county-by-year-by-month level.

Vaccine effectiveness and take-up: We obtain measures of vaccine effectiveness by influenza season and age group 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.[11] These studies measure vaccine effectiveness as the vaccination-induced

[6] We use the Clinical Classifications Software (CCS) from the Agency for Healthcare Research and Quality (AHRQ) to classify relevant influenza ICD codes. These are all 5-digit ICD codes grouped under the following 3-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 3-digit ICD-10-CM codes: J09, J10, J11.

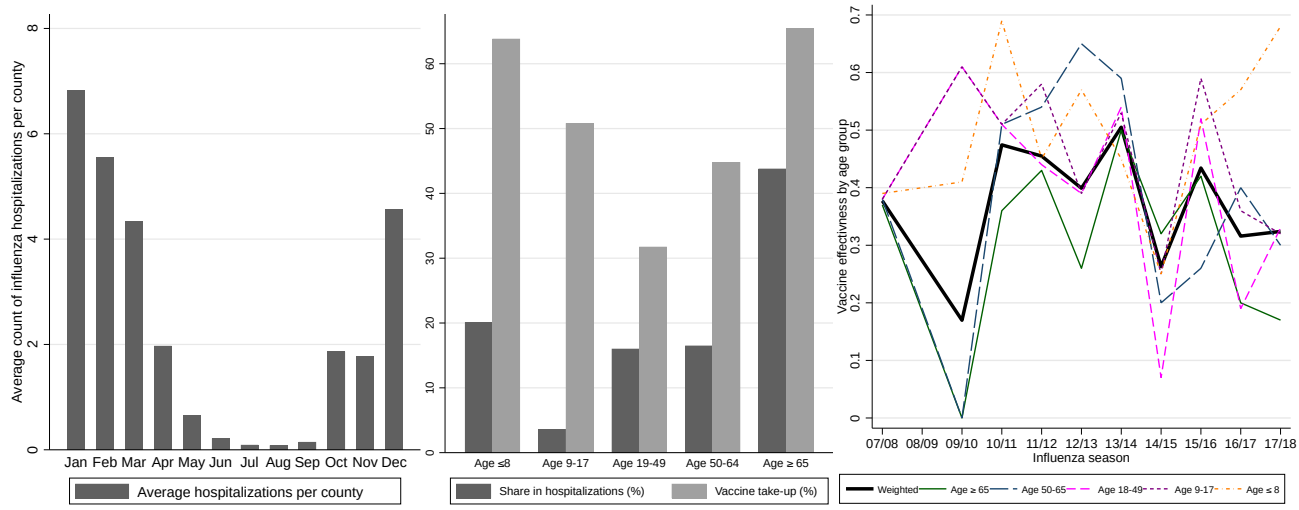
[7] We only do this for counties and year-months in states that report data in that given year.

[8] These are all 5-digit ICD codes grouped under the following 3-digit ICD-9-CM codes: 715, V134; and the following 3-digit ICD-10-CM code: M15, M16, M17, M18, M19.

[9] The EPA pre-aggregates 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% for the main analysis, and show robust results to both data cleaning choices in Appendix A.2.

[10] The AQI captures pollution from particulate matter (PM2.5 or PM10), sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen dioxide (NO_2) and ozone (O_3). See Appendix A.1 for descriptive statistics. The EPA provides further details on AQI calculation in EPA (2018).

[11] 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 $y+1$.



(a) Seasonality of hospitalizations (b) Hospitalizations and vaccination by age (c) Age-specific vaccine effectiveness

Figure 1: Descriptive figures on influenza inpatient hospitalizations and vaccine take-up and effectiveness

Notes: The left Panel (a) shows the average count of influenza inpatient hospitalizations per county-month in the HCUP (2018*b*) data. The middle 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. The right Panel (c) plots (raw) vaccine effectiveness for each age group over influenza seasons (with the exception of 08/09 where no data is available). The thick black line plots our weighted overall vaccine effectiveness.

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.[12]

Figure 1c plots age-specific vaccine effectiveness against influenza seasons, showing variation both across seasons and across 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 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.

Vaccine effectiveness for two age groups, 65 years and older and 8 years and younger, is particularly important. Figure 1b shows both hospitalization incidence and vaccination rates are highest for these two age groups.[13] To generate an *overall* measure of vaccine effectiveness, we weight age-specific vaccine effectiveness (Figure 1c) by both vaccination rates by age groups and shares of influenza hospitalizations by age groups (Figure 1b). We calculate influenza hospitalization shares directly from our HCUP data, and we use data on vaccination rates by influenza season and age group from Lu et al. (2013), Schiller & Euler (2009), CDC (2009, 2015, 2020). Vaccine effectiveness (VE_s) in influenza season s weights raw vaccine effectiveness (VE_{sa}^{raw}) in season s and age group a by vaccination rates ($\overline{VR_a}$) and hospitalization shares ($\overline{HS_a}$) as follows:

$$VE_s = \frac{1}{\sum_a \left(\overline{VR_a} \times \overline{HS_a}\right)} \sum_a VE_{sa}^{raw} \times \overline{VR_a} \times \overline{HS_a}, \tag{1}$$

[12]The odds ratio is approximately the relative risk due to a small number of influenza positive cases (Zhang & Kai 1998).
[13]We use these age cutoffs because they coincide with the common age cutoffs in vaccine effectiveness studies.

where $\overline{VR_a}$ and $\overline{HS_a}$ are simple averages across influenza seasons s, e.g. $\overline{VR_a} = \frac{1}{S}\sum_s VR_{sa}$, and the first term $\frac{1}{\sum_a(\overline{VR_a} \times \overline{HS_a})}$ 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 final measure. This provides a single overall measure of vaccine effectiveness (VE_s) we use for analysis.[14] Figure 1c shows weighted vaccine effectiveness ranges between 0.17 and 0.51 during our study period.

II. Empirical Strategy

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 a Poisson Pseudo Maximum Likelihood (PPML) count data model[15]:

$$H_{cym} = \exp(\beta AQI_{cym-1} + \boldsymbol{X'_{cym}}\boldsymbol{\delta} + \gamma_{csy} + \mu_{ym} + \epsilon_{cym}). \tag{2}$$

We lag the AQI one month to capture exposure to air pollution before hospital admission. We control for a wide variety of both regional and temporal controls. Our preferred specification includes county-by-season-by-year (γ_{csy}) and year-by-month fixed effects (μ_{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 (γ_{csy}) are effectively county by three-month period fixed effects.[16] While county-by-season-by-year fixed effects capture the bulk of climatic differences across counties, we also include weather controls $\boldsymbol{X_{cym}}$ to address the link 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-season-years.[17]

A central challenge is that actual influenza hospitalizations can differ from our observed measure of diagnosed influenza hospitalizations. For example, there may be differences in diagnostic testing for influenza. Our fixed effects absorb bias from discrepancy between actual and observed hospitalizations as long as the ratio between actual and observed hospitalization varies across county-season-years and/or year-months.[18]

County-by-season-by-year effects γ_{csy} similarly control for differences in unobserved confounders that influence pollution exposure and health outcomes across counties separately for every three months, such as demographics, socio-economic factors, or health care access and protocols. Year-by-month fixed effects control for

[14] For our regressions with age-specific outcomes in Table 2, we only use the raw vaccine effectiveness of the corresponding age groups for constructing our overall measure of vaccine effectiveness.

[15] We estimate all models with a PPML estimator (Correia, Guimarães & Zylkin 2019) and show robustness using Ordinary Least Squares (OLS). 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 under-dispersion in the data (Silva & Tenreyro 2006, 2011).

[16] The county-by-season-by-year fixed effects (γ_{csy}) are equivalent to including county-by-year and county-by-season fixed effects separately.

[17] 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 & 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.

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

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 show models using varied fixed effects.

Several econometric challenges exist in evaluating how the influenza vaccine alters the effect of pollution on large-scale infection rates. Vaccinated individuals may reduce avoidance behavior, 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 is another problem: only the most susceptible might seek out vaccines, or only the most cautious. Instead of using variation in vaccination rates[19], 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.[20] Effectiveness based on antigenic drift is, in principle, orthogonal to both air pollution and unobserved determinants of health. This provides insights into how vaccines affect the pollution-induced spread of influenza and provides a test of the effects of pollution on influenza. If vaccine effectiveness moderates the effect of pollution on influenza, it must be that pollution causes influenza.[21]

To estimate the impact of our vaccine effectiveness measure on the pollution-hospitalization relationship, we include an interaction term $AQI_{cym-1} \times VE_s$:

$$H_{cym} = \exp(\beta_1 AQI_{cym-1} + \beta_2 \left(AQI_{cym-1} \times VE_s\right) + \boldsymbol{X'_{cym}\delta} + \gamma_{csy} + \mu_{ym} + \epsilon_{cym}). \tag{3}$$

While random variation in vaccine effectiveness provides a compelling identification strategy, vaccine *take-up* rates may be more amenable to policy intervention. Fortunately, we can use our estimates on vaccine effectiveness to infer the effect of increased take-up. The key insight is that the CDC bases measures of vaccine effectiveness on the number of vaccinated individuals that nonetheless contract the flu in a given season, (see data section). This is conceptually akin to a measure of take-up for a fully effective vaccine. More formally, we define a population measure of the protective effects of the vaccine, which we will refer to as the "effective vaccine take-up" (EVT), as a measure of raw vaccine take-up rates (VR) adjusted for vaccine effectiveness (VE):

$$EVT_s = VR_s \times VE_s. \tag{4}$$

If 50% of people are vaccinated, but the vaccine is only effective for 30% of them, the effective vaccine take-up (EVT) is the same as when only 30% of people are vaccinated but the vaccine is effective in 50% of cases. Equation (4) shows a 10% increase in vaccine effectiveness (VE) has the same effect on EVT and influenza infection rates as a 10% increase in vaccine take-up rates (VR). Therefore, we convert estimated absolute changes in VE_s into relative changes and, together with β_2, infer the effect of changes in vaccine take-up rates.

We cluster standard errors at the county level to allow for arbitrary heteroskedasticity and serial correlation,

[19] Note vaccination take-up rates enter as time-invariant weights in our construction of vaccine effectiveness only to weight the different (raw) vaccine effectiveness values across age groups. Overall take-up does not affect our measure of vaccine effectiveness.

[20] See also Ward (2014) and White (2019) 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.

[21] We cannot distinguish between whether the vaccine is: (i) reducing the probability 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.

Table 1: The effect of air pollution on severe influenza cases

	Influenza is:					
	Primary ICD code		Any ICD code		Only ICD code	
	(1)	(2)	(3)	(4)	(5)	(6)
AQI	0.0089***	0.039***	0.0095***	0.035***	0.015***	0.045**
	(0.0028)	(0.008)	(0.0027)	(0.0072)	(0.0056)	(0.02)
AQI X Vaccine Effectiveness		-0.088***		-0.075***		-0.095
		(0.024)		(0.022)		(0.059)
Observations	17831	17831	20224	20224	3998	3998
Mean of outcome	4.04	4.04	8.34	8.34	0.13	0.13
Mean of AQI predictor	34.46	34.46	34.46	34.46	34.52	34.52
Mean of vac. eff.	-	0.37	-	0.37	-	0.37

Notes: The dependent variable in Columns (1-2) is the count of inpatient hospital admissions with influenza as 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 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 is not available. Vaccine effectiveness is weighted by average vaccination rates and hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson Pseudo-Maximum Likelihood regression with county-by-season-by-year and year-by-month fixed effects 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 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. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

and show robustness to two-way clustering at the added state-year-month level.

III. Results and Discussion

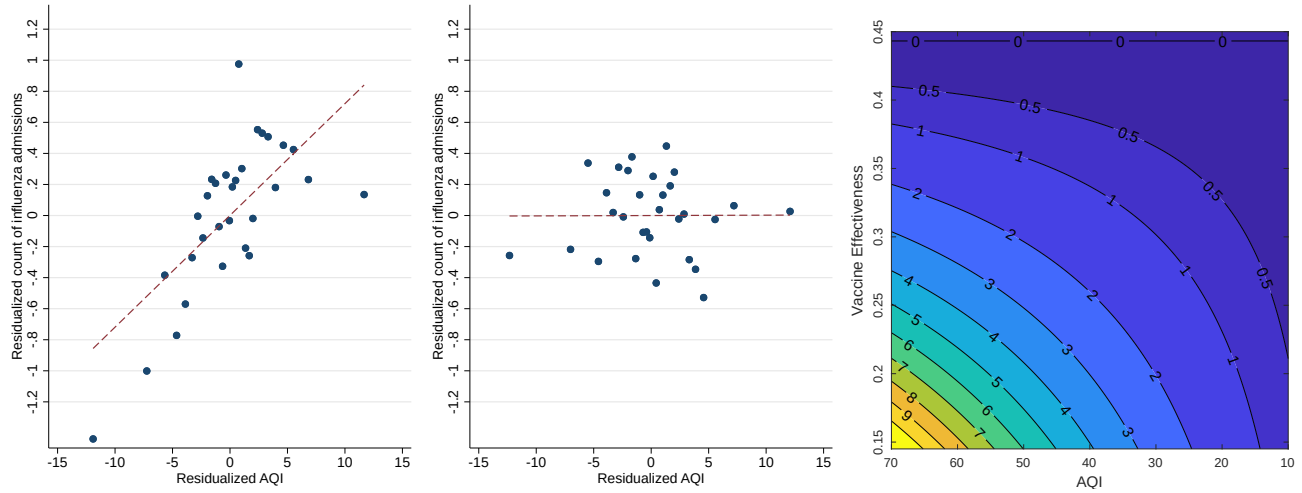
A. Influenza Hospitalizations

Table 1 shows estimates from our PPML regressions. Coefficients represent the AQI semi-elasticity 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. Estimates from Column (1) correspond to Equation (1) and imply a 1-unit increase in the monthly AQI results in a 0.89% increase in influenza inpatient admissions. To put this estimate in national context, a one-standard-deviation increase in AQI (10.9-unit increase in our data) amounts to approximately 7,760 (10.2%) additional inpatient hospitalizations for a 6-month influenza season in the U.S.[22]

To explore the moderating role of vaccine effectiveness, Figure 2 shows the regression-adjusted relationship between AQI and influenza admissions separately in seasons of low vaccine effectiveness in Panel (a) and high vaccine effectiveness in Panel (b), as determined by an effectiveness median (0.38) sample split. The relationship between air quality and admissions rates is positively sloped in Panel (a), indicating the AQI affects flu admissions when the vaccine is ineffective. This relationship flattens completely in seasons of high vaccine effectiveness, shown in Panel (b), suggesting an effective vaccine nullifies the relationship between pollution and the flu. This does not imply a high vaccine effectiveness eliminates all influenza hospitalizations. Rather, a sufficiently high vaccine effectiveness eliminates flu cases attributable to pollution.

To test for the moderating role of vaccine effectiveness, we present estimates of Equation (3) using the con-

[22]We use the 10.9-unit increase and the coefficient 0.0089 for the relative increase $(\exp(0.0089 * 10.9) - 1 = 0.1019$, 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) (Bureau 2018) and by the 6 months within a influenza season. We only count cases with primary diagnosis influenza, making this estimate of absolute numbers a lower bound. Using hospitalization with any influenza diagnosis (shown in Column (3)) more than doubles the additional predicted cases because the base of hospital admissions is much larger.



(a) Low vaccine effectiveness (b) High vaccine effectiveness (c) Additional hosp. charges (bil. $US)

Figure 2: Air quality, vaccine effectiveness and influenza hospitalizations and charges

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. The panels show the relationship for seasons with below (a) and above (b) median vaccine effectiveness. Panel (c) shows a contour plot of additional inpatient hospitalization charges for different weighted vaccine effectiveness and AQI levels. The contour lines indicate the additional charges in billion $US aggregated across the U.S. per influenza season from October to March. They are calculated from our PPML estimates in Column 2 of Table 1, the average hospital charges per county-year-month (117 th. $US), the count of U.S. county equivalents (3142) and months in an influenza season (6). The charges are additional compared to a zero average AQI, conditional on our controls and fixed effects.

tinuous measure of vaccine effectiveness in Column (2) of Table 1. Vaccine effectiveness substantially affects pollution driven influenza cases. Our negative interaction coefficient in Column (2) implies a weighted vaccine effectiveness of 44%, close to the maximum of 51% in our sample, is sufficient to nullify the link between air pollution and influenza hospitalizations.[23] Due to the equivalence of the relative effects of vaccine effectiveness and take-up rates, this implies that, based on an average vaccine effectiveness of 37%, a 19% increase in vaccine take-up would cause the full neutralizing effect. The 25th percentile of vaccine effectiveness (26%) would require a 69% increase in vaccine take-up to neutralize pollution effects. For a vaccine effectiveness over the 75th percentile, the effects of pollution are completely nullified for an average level of vaccine take-up.

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 over-counting 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 but count patients that have any (primary or secondary) influenza diagnosis. The average number of influenza admissions per county-year-month is roughly double (8.34) compared with our baseline approach (4.04). The estimated coefficients, however, 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 only counting hospital admissions where the only diagnosis is influenza. This reduces the average count of admissions per county-year-month to 0.13 (the majority of influenza hospital admissions have further

[23] There is 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 & Goldfarb 2007, Ward 2014).

Table 2: Heterogeneity by age and race

	≤ 17y		18-64y		≥ 65y		Black/Hispanic		White	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
AQI	0.0094***	0.02*	0.011***	0.026***	0.0054**	0.025***	0.01*	0.049***	0.01***	0.037***
	(0.003)	(0.012)	(0.0037)	(0.0068)	(0.0027)	(0.0057)	(0.0056)	(0.015)	(0.0024)	(0.0075)
AQI X Vaccine Effectiveness		-0.023		-0.04**		-0.068***		-0.11***		-0.079***
		(0.024)		(0.019)		(0.021)		(0.04)		(0.023)
Observations	11428	11428	13752	13752	13746	13746	7819	7819	15699	15699
Mean of outcome	0.89	0.89	1.33	1.33	1.82	1.82	0.94	0.94	2.66	2.66
Mean of AQI predictor	34.46	34.46	34.46	34.46	34.46	34.46	34.46	34.46	34.46	34.46
Mean of vac. eff.	-	0.48	-	0.4	-	0.3	-	0.37	-	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 or race 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 is not available. Vaccine effectiveness is weighted by average vaccination rates and hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson Pseudo-Maximum Likelihood regression with county-by-season-by-year and year-by-month fixed effects 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 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. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

influenza-induced complications, e.g., pneumonia). The estimated coefficients are, however, also comparable to our baseline estimates, though with larger standard errors given the considerable drop in sample size.

We explore heterogeneity by age and race in Table 2. Columns 1 through 6 show results for three distinct age groups: children and adolescents ($\leq$ 17 years), adults under age 65, and adults over age 65. Patterns across groups are similar, with overlapping confidence intervals on our estimates. The interaction with vaccine effectiveness is largest in magnitude for the 65 years and older group. As Figure 1b shows, this group also has the highest vaccination rate. A higher vaccine take-up driving a larger direct impact of vaccine effectiveness is consistent with the larger coefficient on the interaction for this group.[24]

Estimates are similar across racial and ethnic groups (Blacks/Hispanics and Whites in Columns 7 through 10), with overlapping confidence intervals. The magnitude of the protective effect of the vaccine is slightly higher for Blacks/Hispanics. Given racial and ethnic differences in pollution exposure (Banzhaf, Ma & Timmins 2019, Colmer et al. 2020, Currie, Voorheis & Walker 2020), this aligns with vaccines yielding greater returns for those in more polluted areas, where additional pollution may generate disproportionate health harms and vaccines disproportionate benefits accordingly. Higher levels of pollution exposure amongst Black and Hispanic communities could also help explain historically higher influenza burdens experienced by those communities (e.g. Quinn et al. 2011). Our results suggest air quality control could be an additional policy lever to help reduce severe influenza cases among highly affected groups.

Table 3 explores robustness of our main results. In Columns 1 and 2, we replace our county-by-season-by-year fixed effects with coarser county-by-influenza season effects. Data are on a calendar year basis, so these coarser fixed effects capture fewer discrepancies across reporting years, which introduces error and appears to attenuates our main results. Including the VE interaction term causes all terms to be statistically significant again, further highlighting the benefit of this additional source of variation. In Columns 3 and 4 we instead add state-by-month (e.g., New York in February) fixed effects to our baseline to allow seasonality to vary by state. The coefficient on the level effect of AQI is smaller than in our baseline and no longer statistically significant, but the specification

[24]Since vaccines have private but also external benefits, vaccine take-up of any one group generates positive spillovers to other groups.

Table 3: The effect of air pollution on severe influenza cases: robustness and falsification

	Fewer FE		More FE		Two-way clus. SE		Osteoarthritis	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AQI	0.00051	0.025***	0.0024	0.022***	0.0089**	0.039***	-0.00046	-0.000055
	(0.0036)	(0.0089)	(0.0024)	(0.0069)	(0.0044)	(0.013)	(0.00032)	(0.0008)
AQI X Vaccine Effectiveness		-0.069***		-0.055***		-0.088**		-0.001
		(0.026)		(0.02)		(0.035)		(0.0023)
Observations	21482	21482	17831	17831	17831	17831	24562	24562
Mean of outcome	4.04	4.04	4.04	4.04	4.04	4.04	40.06	40.06
Mean of AQI predictor	34.46	34.46	34.46	34.46	34.46	34.46	34.46	34.46
Mean of vac. eff.	-	0.37	-	0.37	-	0.37	-	0.37
FE County X Inf. Seas.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE County X Year	No	No	Yes	Yes	Yes	Yes	Yes	Yes
FE State X Month	No	No	Yes	Yes	No	No	No	No
FE Month X Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in Columns (1-6) is the count of inpatient hospital admissions with influenza as primary diagnosis within a county-year-month. The dependent variable in Columns (7-8) is the count of inpatient hospital admissions with osteoarthritis as primary diagnosis within a county-year-month. Columns (1-2) and (3-4) have different sets of fixed effects than our baseline which contains county-by-season-by-year and year-by-month fixed effects. Columns (5-6) contains standard errors in parentheses that are two-way clustered at the county level and the state-by-year-by-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 is not available. Vaccine effectiveness is weighted by average vaccination rates and hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson Pseudo-Maximum Likelihood regression with included fixed effects county-by-season-by-year as indicated 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 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 except for Columns (5-6). *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

with the interaction remains close to our baseline estimates. In Columns 5 and 6 we calculate standard errors by two-way clustering on county and state-by-year-by-month to allow for more spatial correlation. This only slightly increases our standard errors. Finally, to further bolster our causal claims, we perform a falsification test by repeating our analysis using osteoarthritis as a health outcome, which should not correlate with air quality, influenza, or vaccine effectiveness. Both Column (7) and Column (8) indicate precise zero coefficients on the effect of AQI and the interaction with vaccine effectiveness.

Appendix A.2 includes further robustness checks: (i) including off-season (Apr-Sep) county-year-months with positive influenza hospitalizations, (ii) using missing values instead of zeros for county-year-months with no hospital admissions, (iii) using OLS instead of the PPML estimator, (iv) using different or no weather controls, and (v) different winsorization or interpolation of the raw AQI data. All have little effect on our estimates. We also show the effects on emergency department outpatient hospitalizations, which are more frequent but also less severe, are similar to those for inpatient hospitalizations.

B. Medical Charges and Policy Implications

We next calculate the additional hospital charges attributable to influenza to get a sense of the costs generated by air pollution.[25] We use our estimates from Column 1 of Table 1, together with the average charges per county-year-month ($US 117 thousand), to draw a contour plot of additional hospital charges spanning the support of AQI and vaccine effectiveness in our data, shown in Figure 2c. Contour lines show pollution-induced influenza inpatient hospitalization charges at various levels of AQI (decreasing along the horizontal axis so as to represent

[25]Hospital charges are around $US 29 thousand per patient per influenza diagnosed inpatient hospitalization, so $US 117 thousand per county-year-month. Note hospital charges are distinct from hospital costs, which are notoriously difficult to ascertain because they differ significantly across institutions and units within institutions. Further, these estimates ignore indirect costs to patients, such as forgone earnings.

an improvement in air quality) and vaccine effectiveness (increasing along the vertical axis) during an influenza seasons across the U.S. in billions of $US. Contour lines are similar to isocost curves, but instead of measuring levels of charges, they represent *additional* charges compared to an AQI of zero.

Figure 2c illustrates our main results in terms of additional hospital charges. First, at the top of the figure, when vaccine effectiveness is high, an increase in AQI – no matter how large – has no impact on flu hospitalizations charges because of the protective nature of the vaccine.[26] In contrast, at the bottom of the figure when vaccine effectiveness is low, even small changes in the AQI generate large increases in additional influenza-specific hospitalization charges. Going from an AQI of 40 to 50 generates roughly 2 billion $US in additional influenza inpatient hospitalization charges at a vaccine effectiveness of 0.15. Second, at the right of the figure, for good air quality (low AQI), a drop in vaccine effectiveness generates little additional pollution-driven influenza hospitalization charges (though influenza cases that are not pollution driven still might be greatly affected). A drop from 0.4 to 0.2 effectiveness only generates around 0.5 billion $US in additional pollution-driven influenza charges. On the left of the figure for high AQI values, however, the same drop in vaccine effectiveness from 0.4 to 0.2 generates around 7 billion $US in additional pollution-induced influenza hospital charges.

Are vaccine and air quality policies substitutes or complements in preventing pollution-induced influenza cases? The answer depends on the stochastic nature of the two health shocks. From any given point in the space of 2c, the ex-ante marginal benefit from improving vaccine effectiveness or air quality decreases in the level of the other variable. If both cleaner air quality and vaccine effectiveness were deterministic policy outcomes, they would serve as substitutes. Vaccine effectiveness, however, is a stochastic outcome due to unforeseen and random antigenic drift and high variability from season to season (see Figure 1c). Air quality is also inherently stochastic because of imperfect control of emissions, variations in activities that cause emissions, the role weather plays in converting emissions to pollution, and natural sources of emissions, such as wildfires. Random variations in both vaccine effectiveness and air quality results in a higher ex-post marginal benefit of the other variable. The stochastic nature of both factors thus introduces complementarities between pollution control and vaccination policy.

For seasons with poor vaccine effectiveness, improved air quality can provide an important hedge to reduce influenza cases. Similarly, for seasons with higher local air pollution, effective vaccines or higher vaccine take-up rates (see Equation (4)) can provide protective effects from pollution-driven influenza. A back of the envelope calculation suggests that a 10% (3.5 AQI points) reduction in the AQI in an historically bad vaccine year (17% effectiveness) would avert 6,146 (8.2%) hospitalizations across the U.S. or $US 178 million in influenza medical charges, while a 10% 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 8,908 (13.2%) of hospitalizations, or $US 292 million.

IV. Conclusion

Using a rich longitudinal dataset, we provide evidence air pollution increases seasonal influenza hospitalization rates, and that an effective influenza vaccine greatly diminishes this relationship. Our empirical strategy,

[26] This does not imply higher AQI has no impact on hospitalizations costs of any cause, but that is has no impact on influenza hospitalizations.

based on the stochastic nature of vaccine effectiveness across influenza seasons, limits risks of confounding. We highlight how improving air quality and increasing vaccination rates can both yield substantial social returns for fighting influenza. This is especially important in dense urban centers around the world and developing countries in particular, where pollution and vaccination externalities are likely highest (de Lataillade, Auvergne & Delannoy 2009). Pollution controls provide an important hedge against the regular shocks to vaccine effectiveness, while increasing vaccine uptake can hedge against the stochastic relationship between emissions and ambient concentrations of pollution arising from complex atmospheric chemistry and external events such as changes in prevailing wind patterns, extreme heat, and wildfires. When these policies work in conjunction, they help reduce medical spending, avoid lost productivity, and reduce loss of life.

Our insights regarding compounding risks from pollution and flu may extend to other viral respiratory illnesses with similar etiological pathways, including the current COVID-19 pandemic.[27] Though research remains preliminary, evidence suggests significant positive correlations between COVID-19 hospitalizations and pollution levels (Wu et al. 2020). Since large scale reductions in economic activity aimed at reducing viral spread have reduced current air pollution (NASA 2020), the importance of this relationship may be masked in the data, even if the pollution-COVID-19 link is causal. As economic activity resumes, pollution will increase, which may compound the threat from COVID-19 infections. If governments suspend environmental regulations in an effort to bolster the economic recovery, as has been recently seen in the U.S. (Bodine 2020), hospitalizations and deaths from the pandemic may be further hastened. Absent a highly effective vaccine with widespread take-up, our results suggest a different policy direction, where additional environmental controls serve as a complementary investment to optimally manage the full harms from this current global health threat as well as potential future pandemics.

REFERENCES

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): 6558–6563.

Bodine, Susan Parker. 2020. *Memorandum: COVID-19 Implications for EPA's Enforcement and Compliance Assurance Program.* United States Environmental Protection Agency.

Boulier, Bryan L, Tejwant S Datta, and Robert S Goldfarb. 2007. "Vaccination externalities." *The BE Journal of Economic Analysis & Policy*, 7(1).

Brauer, Michael, Gerard Hoek, Patricia Van Vliet, Kees Meliefste, Paul H Fischer, Alet Wijga, Laurens P Koopman, Herman J Neijens, Jorrit Gerritsen, Marjan Kerkhof, 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): 1092–1098.

Bureau, United States Census. 2018. *2018 FIPS Codes.* United States Census Bureau.

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." *Appl. Environ. Microbiol.*, 76(9): 2712–2717.

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, National Health Interview Survey (NHIS), United States.* Centers for Disease Control and Prevention.

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.

[27] See e.g. Cui et al. (2003) for evidence on SARS-CoV.

CDC. 2019. *Seasonal Flu Vaccine Effectiveness Studies.* Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases (NCIRD).

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.

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): 1211–1216.

Ciencewicki, Jonathan, and Ilona Jaspers. 2007. "Air pollution and respiratory viral infection." *Inhalation toxicology*, 19(14): 1135–1146.

Clay, Karen, Joshua Lewis, and Edson Severnini. 2018. "Pollution, infectious disease, and mortality: evidence from the 1918 Spanish influenza pandemic." *The Journal of Economic History*, 78(4): 1179–1209.

Cohen, Aaron J, Michael Brauer, Richard Burnett, H Ross Anderson, Joseph Frostad, Kara Estep, Kalpana Balakrishnan, Bert Brunekreef, Lalit Dandona, Rakhi Dandona, 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): 1907–1918.

Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis. 2020. "Disparities in PM2. 5 air pollution in the United States." *Science*, 369(6503): 575–578.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2019. "PPMLHDFE: Fast poisson estimation with high-dimensional fixed effects." *arXiv preprint: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): 15.

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*.

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.

Diamond, Gill, Diana Legarda, and Lisa K Ryan. 2000. "The innate immune response of the respiratory epithelium." *Immunological reviews*, 173: 27–38.

EPA. 2018. *Technical Assistance Document for the Reporting of Daily Air Quality.* United States Environmental Protection Agency.

EPA. 2020. *Air Quality System Data Mart.* US Environmental Protection Agency.

Flannery, Brendan, Jessie R Chung, Arnold S Monto, Emily T Martin, Edward A Belongia, Huong Q McLean, Manjusha Gaglani, Kempapura Murthy, Richard K Zimmerman, Mary Patricia Nowalk, 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, Michael L Jackson, Lisa A Jackson, Arnold S Monto, et al. 2020. "Spread of antigenically drifted influenza A (H3N2) viruses and vaccine effectiveness in the United States during the 2018–2019 season." *The 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): 572–578.

Gaglani, Manjusha, Jessica Pruszynski, Kempapura Murthy, Lydia Clipper, Anne Robertson, Michael Reis, Jessie R Chung, Pedro A Piedra, Vasanthi Avadhanula, Mary Patricia Nowalk, et al. 2016. "Influenza vaccine effectiveness against 2009 pandemic influenza A (H1N1) virus differed by vaccine type during 2013–2014 in the United States." *The Journal of infectious diseases*, 213(10): 1546–1556.

Gourieroux, Christian, Alain Monfort, Alain Trognon, et al. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica*, 52(3): 701–720.

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): 430–438.

Griffin, Marie R, Arnold S Monto, Edward A Belongia, John J Treanor, Qingxia Chen, Jufu Chen, H Keipp Talbot, Suzanne E Ohmit, Laura A Coleman, Gerry Lofthus, 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).

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): 405–417.

Harper, GJ. 1961. "Airborne micro-organisms: survival tests with four viruses." *Epidemiology & Infection*, 59(4): 479–486.

HCUP. 2018*a*. *HCUP State Emergency Department Databases (SEDD)*. Healthcare Cost and Utilization Project, Agency for Healthcare Research and Quality, Rockville, MD.

HCUP. 2018*b*. *HCUP State Inpatient Databases (SID)*. Healthcare Cost and Utilization Project, Agency for Healthcare Research and Quality, Rockville, MD.

Ijaz, M Khalid, Syed A Sattar, C Margaret Johnson-Lussenburg, and VS Springthorpe. 1985. "Comparison of the airborne survival of calf rotavirus and poliovirus type 1 (Sabin) aerosolized as a mixture." *Appl. Environ. Microbiol.*, 49(2): 289–293.

Iuliano, A Danielle, Katherine M Roguski, Howard H Chang, David J Muscatello, Rakhee Palekar, Stefano Tempia, Cheryl Cohen, Jon Michael Gran, Dena Schanzer, Benjamin J Cowling, et al. 2018. "Estimates of global seasonal influenza-associated respiratory mortality: a modelling study." *The 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, Edward A Belongia, Huong Q McLean, Manjusha Gaglani, et al. 2017. "Influenza vaccine effectiveness in the United States during the 2015–2016 season." *New England Journal of Medicine*, 377(6): 534–543.

Jaspers, Ilona, Jonathan M Ciencewicki, 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 LC Marr. 2015. "Simulation of vertical concentration gradient of influenza viruses in dust resuspended by walking." *Indoor Air*, 25(4): 428–440.

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): 2036–2044.

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 2.5 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 2.5 and PM 10 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 Pathog*, 3(10): e151.

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): 1478–1487.

McLean, Huong Q, Mark G Thompson, Maria E Sundaram, Burney A Kieke, Manjusha Gaglani, Kempapura Murthy, Pedro A Piedra, Richard K Zimmerman, Mary Patricia Nowalk, Jonathan M Raviotta, et al. 2015. "Influenza vaccine effectiveness in the United States during 2012–2013: variable protection by age and virus type." *The Journal of infectious diseases*, 211(10): 1529–1540.

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 (GES DISC),Greenbelt, Maryland, USA.

NASA. 2020. *Airborne Nitrogen Dioxide Plummets Over China.* NASA Earth Observatory.

Ohmit, Suzanne E, Mark G Thompson, Joshua G Petrie, Swathi N Thaker, Michael L Jackson, Edward A Belongia, Richard K Zimmerman, Manjusha Gaglani, Lois Lamerato, Sarah M Spencer, 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): 319–327.

Putri, Wayan CWS, David J Muscatello, Melissa S Stockwell, and Anthony T Newall. 2018. "Economic burden of seasonal influenza in the United States." *Vaccine*, 36(27): 3960–3966.

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): 285–293.

Rivas-Santiago, César E, Srijata Sarkar, Pasquale Cantarella, Álvaro Osornio-Vargas, Raúl Quintana-Belmares, Qingyu Meng, Thomas J Kirn, Pamela Ohman Strickland, Judith C Chow, John G Watson, et al. 2015. "Air pollution particulate matter alters antimycobacterial respiratory epithelium innate immunity." *Infection and immunity*, 83(6): 2507–2517.

Rolfes, Melissa A, Brendan Flannery, Jessie R Chung, Alissa O'Halloran, Shikha Garg, Edward A Belongia, Manjusha Gaglani, Richard K Zimmerman, Michael L Jackson, Arnold S Monto, et al. 2019. "Effects of influenza vaccination in the United States during the 2017–2018 influenza season." *Clinical Infectious Diseases*, 69(11): 1845–1853.

Schiller, Jeannine S, and GL Euler. 2009. "Vaccination coverage estimates from the National Health Interview Survey: United States, 2008." *National Center for Health Statistics (NCHS). Health E-Stat. Centers for Disease Control and Prevention (CDC)*.

Shaman, Jeffrey, and Melvin Kohn. 2009. "Absolute humidity modulates influenza survival, transmission, and seasonality." *Proceedings of the National Academy of Sciences*, 106(9): 3243–3248.

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 Biol*, 8(2): e1000316.

Silva, JMC Santos, and Silvana Tenreyro. 2006. "The log of gravity." *The Review of Economics and statistics*, 88(4): 641–658.

Silva, JMC Santos, and Silvana Tenreyro. 2011. "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator." *Economics Letters*, 112(2): 220–222.

Tellier, Raymond. 2009. "Aerosol transmission of influenza A virus: a review of new studies." *Journal of the Royal Society Interface*, 6(suppl_6): S783–S790.

Treanor, John J, H Keipp Talbot, Suzanne E Ohmit, Laura A Coleman, Mark G Thompson, Po-Yung Cheng, Joshua G Petrie, Geraldine Lofthus, Jennifer K Meece, John V Williams, 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): 951–959.

Tschofen, Peter, Inês L Azevedo, and Nicholas Z Muller. 2019. "Fine particulate matter damages and value added in the US economy." *Proceedings of the National Academy of Sciences*, 116(40): 19857–19862.

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. 2019. "Measuring Social and Externality Benefits of Influenza Vaccination." *Journal of Human Resources*.

Wolkoff, Peder. 2018. "Indoor air humidity, air quality, and health–An overview." *International journal of hygiene and environmental health*, 221(3): 376–390.

Wong, Chit Ming, Lin Yang, Thuan Quoc Thach, Patsy Yuen Kwan Chau, King Pan Chan, G Neil Thomas, Tai Hing Lam, Tze Wai Wong, Anthony J Hedley, and JS Malik Peiris. 2009. "Modification by influenza on health effects of air pollution in Hong Kong." *Environmental health perspectives*, 117(2): 248–253.

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 preprint: 10.1101/2020.04.05.20054502*, , (10.1101/2020.04.05.20054502).

Xia, Youlong, Kenneth Mitchell, Michael Ek, Justin Sheffield, Brian Cosgrove, Eric Wood, Lifeng Luo, Charles Alonge, Helin Wei, Jesse Meng, 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 F Yu Kai. 1998. "What's the relative risk?: A method of correcting the odds ratio in cohort studies of common outcomes." *Jama*, 280(19): 1690–1691.

Zimmerman, Richard K, Mary Patricia Nowalk, Jessie Chung, Michael L Jackson, Lisa A Jackson, Joshua G Petrie, Arnold S Monto, Huong Q McLean, Edward A Belongia, Manjusha Gaglani, et al. 2016. "2014–2015 Influenza Vaccine Effectiveness in the United States by Vaccine Type." *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 63(12): 1564.

APPENDIX

When Externalities Collide: Influenza and Pollution

by Joshua Graff Zivin[1,2,*], Matthew Neidell[3,2,*], Nicholas J. Sanders[4,2,*], Gregor Singer[1,*]

[1]University of California, San Diego

[2]National Bureau of Economic Research

[3]Columbia University

[4]Cornell University

[*]Correspondence to: Joshua Graff Zivin: jgraffzivin@ucsd.edu, Matthew Neidell: mn2191@cumc.columbia.edu, Nicholas J. Sanders: njsanders@cornell.edu, Gregor Singer: gsinger@ucsd.edu

A.1 Additional Descriptive Statistics

Table A.1 contains states and years with available admission months and patient zip codes in the HCUP (2018*b*) inpatient hospitalization data we use. Table A.2 contains summary statistics at the county-year-month level for inpatient hospital admissions with a primary influenza diagnosis, associated hospital charges, and the average monthly AQI. We use the standard deviation of the AQI during the influenza season (10.9), the average inpatient hospitalization admissions (4.04) and charges (117,000 US$) for the calculation of absolute effects based on our Poisson Pseudo-Maximum Likelihood estimates.

To further illustrate the influenza seasonality, we use data on the timing of national influenza-like illnesses from the Centers for Disease Control and Prevention (CDC 2020). Figure A.1 shows that the seasonality of inpatient hospitalizations in our data matches closely with general influenza-like illnesses reported by the CDC.

The AQI is based on multiple pollutants, but for each county-day, a single pollutant is the defining pollutant of the AQI (EPA 2018). Figure A.2 shows which pollutants are the main defining pollutants of the AQI during the influenza season from October through March for three different intervals covering our sample. Particulate matter (PM2.5 and PM10) and ozone are the defining pollutants in the AQI for the majority of cases in each time period.

Table A.1: Data coverage with available zip codes and admission months

Arizona	2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017
Arkansas	2009
Colorado	2007,2008,2009,2010,2011,2012
Hawaii	2009
Iowa	2009
Kentucky	2007,2008,2009,2010,2011,2012,2013,2014
Maryland	2009,2010,2011,2012
Massachusetts	2007,2008,2009,2010,2011,2012,2013,2014
Michigan	2008,2009,2010,2011,2012,2013,2014,2015,2016,2017
Minnesota	2014,2015,2016
Nevada	2010,2011,2012,2013,2014,2015
New Jersey	2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017
New York	2007,2008,2009,2010,2011,2012,2013,2014,2015
North Carolina	2008,2009,2010,2011,2012,2013,2014,2015,2016,2017
Oregon	2008,2009
Rhode Island	2007,2008,2009,2010,2011,2012,2013,2014,2015
South Dakota	2009
Utah	2009
Vermont	2009
Washington	2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017
Wisconsin	2009

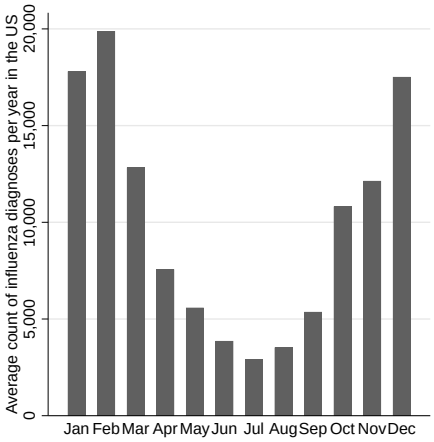
Notes: The table shows the states and years with available admission month and patient zip code used in the analysis for influenza hospitalizations.

Table A.2: Summary statistics of influenza hospitalizations and air pollution (AQI)

		Mean	SD	Min	5th p.	10th p.	25th p.	75th p.	90th p.	95th p.	Max
Hospital admissions per county per month	Oct-Mar	4.04	16.3	0	0	0	0	2	8	17	588
	Apr-Sep	0.526	3.41	0	0	0	0	0	1	2	170
Hospital charges (th. USD) per county per month	Oct-Mar	117	567	0	0	0	0	39.1	202	503	23729
	Apr-Sep	16.7	124	0	0	0	0	0	18	57.5	6883
Average AQI across county-months	Oct-Mar	34.5	10.9	7.14	16.3	21	28	40.6	47.3	52.9	72.4
	Apr-Sep	42.9	14.1	11.3	17.8	23.5	35.2	50.2	59.7	67.6	84.8

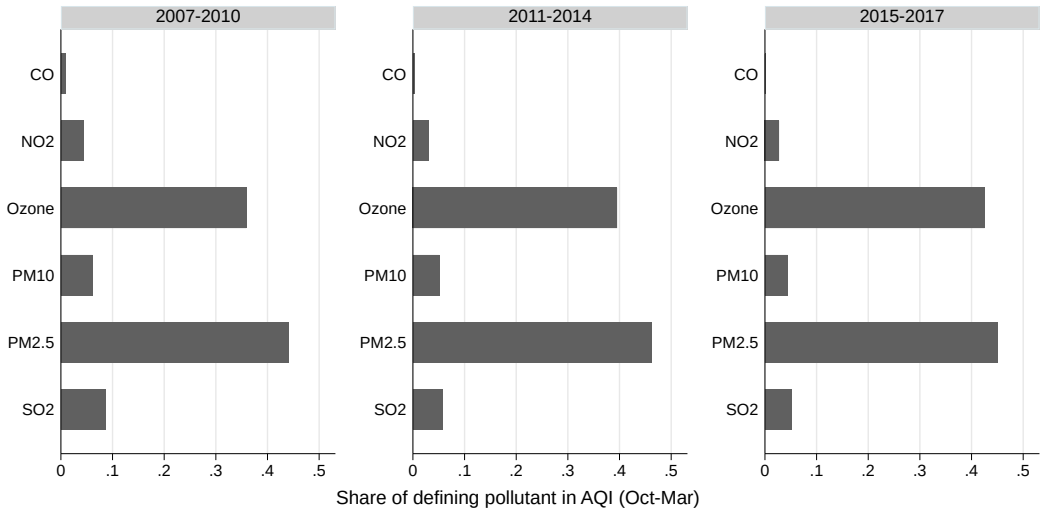
Notes: The table shows summary statistics for influenza diagnosed inpatient hospital admissions and charges, and air pollution measured by the AQI. We pool and report data separately by the influenza season of October through March and the off season of April through September. The AQI statistics are based on the coverage of the hospitalization sample.

Figure A.1: Influenza-like illnesses in U.S.



Notes: The figure shows the distribution of recorded influenza-like illnesses from CDC (2020), which includes non-hospitalized cases. Data are pooled across the U.S. spanning 1997-2019. Not all health providers report to the Influenza-Like Illness (ILI) Network, and the number of providers reporting grew over time so total number of cases is a lower bound of true infection rates.

Figure A.2: Defining pollutants of the AQI



Notes: The figure shows each pollutant's share in days when it was the defining pollutant for calculating the AQI at the county-day level. The shares in days are calculated for the three to four year periods as indicated and are based on the months of the influenza season (Oct-Mar). The data on defining pollutants comes from EPA (2020).

A.2 Additional Robustness Checks

In this section we provide additional robustness checks for our main results. In Columns 1 and 2 of Table A.3, we add all county-by-year-by-month cells which are non-zero valued, i.e. we include off-seasonal cells with influenza cases. In Columns 3 and 4 of Table A.3 we instead drop all zero-valued county-by-year-by-month cells during influenza season, which is reflected in the higher mean of the outcome. Both these results are consistent with the results in the main paper.

In Columns 5 and 6 of Table A.3, we estimate an ordinary least squares (OLS) model instead of a Poisson Pseudo Maximum Likelihood (PPML) model. The patterns are the same as the results in the main paper, and the coefficients can be interpreted as a level effect instead of a semi-elasticity.

In Columns 7 and 8 of Table A.3, we use the data on outpatient hospitalizations at emergency departments instead of inpatient hospitalizations. The mean of the outcome (34.2) is higher than our baseline (4.04), reflecting that outpatient hospitalizations are are more common than inpatient hospitalizations, but also less severe cases. The semi-elasticities are similar to our estimates in the main paper.

In Columns 1 to 4 of Table A.4, we use different weather controls. In Columns 1 and 2, we drop all weather controls. In Columns 3 and 4 we use a second degree polynomial in temperature and specific humidity with all interactions and linear terms for wind speed and precipitation. In Columns 5 to 8 we use different steps to construct our AQI variable. In Columns 5 and 6, we do not winsorize the top and bottom 1 percent of the AQI. In Columns 7 and 8, we do not use the average of adjacent county readings for missing AQI values at the county-by-year-by-month level. Our main results are robust to all of these checks.

REFERENCES

CDC (2020), *U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet)*, Centers for Disease Control and Prevention.

EPA (2018), *Technical Assistance Document for the Reporting of Daily Air Quality*, United States Environmental Protection Agency.

EPA (2020), *Air Quality System Data Mart*, US Environmental Protection Agency.

HCUP (2018), *HCUP State Inpatient Databases (SID)*, Healthcare Cost and Utilization Project, Agency for Healthcare Research and Quality, Rockville, MD.

Table A.3: Off-seasonal cases, removing zeros in outcome, OLS and outpatient hospitalizations

	Incl. off-seas. cases		Removing zeros		OLS		Outpatient hospitalizations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AQI	0.0081***	0.033***	0.0091***	0.041***	0.041**	0.12**	0.017***	0.049***
	(0.0023)	(0.0062)	(0.0029)	(0.0082)	(0.016)	(0.051)	(0.0033)	(0.0097)
AQI X Vaccine Effectiveness		-0.068***		-0.094***		-0.2*		-0.098***
		(0.016)		(0.025)		(0.11)		(0.031)
Observations	21860	21860	9761	9761	24742	24742	10099	10099
Mean of outcome	3.97	3.97	8.76	8.76	4.04	4.04	34.2	34.2
Mean of AQI predictor	35.68	35.68	35.81	35.81	34.46	34.46	35.13	35.13
Mean of vac. eff.	-	0.37	-	0.36	-	0.37	-	0.37

Notes: The dependent variable in Columns (1-6) is the count of inpatient hospital admissions with influenza as primary diagnosis within a county-year-month. The dependent variable in Columns (7-8) is the count of outpatient hospital admissions (emergency departments) with influenza as any (primary or secondary) 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 is not available. In Columns (1-2), we add all off-seasonal county-year-month cells with non-zero admissions. In Columns (3-4) we remove all zero-valued county-year-months cells during the influenza season. Vaccine effectiveness is weighted by average vaccination rates and hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson Pseudo-Maximum Likelihood regression in Columns (1-4,7-8) and from an OLS regression in Columns (5-6), both with county-by-season-by-year and year-by-month fixed effects 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 number of included observations can vary across different columns due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table A.4: Alternative weather controls, not winsorizing or interpolating the AQI

	No weather contr.		Polyn. weather contr.		AQI not winsorized		AQI not interpolated	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AQI	0.0076***	0.027***	0.0072***	0.033***	0.0072***	0.033***	0.0083***	0.035***
	(0.0026)	(0.0084)	(0.0025)	(0.0078)	(0.0025)	(0.0081)	(0.003)	(0.009)
AQI X Vaccine Effectiveness		-0.057**		-0.076***		-0.075***		-0.079***
		(0.024)		(0.022)		(0.023)		(0.027)
Observations	18062	18062	17831	17831	17831	17831	9042	9042
Mean of outcome	4.04	4.04	4.04	4.04	4.04	4.04	4.04	4.04
Mean of AQI predictor	34.46	34.46	34.46	34.46	34.62	34.62	35.35	35.35
Mean of vac. eff.	-	0.37	-	0.37	-	0.37	-	0.37

Notes: The dependent variable is the count of inpatient hospital admissions 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 is not available. Vaccine effectiveness is weighted by average vaccination rates and hospitalization shares across age groups and is measured between 0 (low) and 1 (high). The results are from a Poisson Pseudo-Maximum Likelihood regression with county-by-season-by-year and year-by-month fixed effects as well as weather controls. In Columns (1-4) weather controls are included as described. In Columns (5-8) 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 number of included observations can vary across different columns due to fixed effects and varied counts in each county-year-month cell. Standard errors in parentheses are clustered at the county level. *** Significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.