NBER WORKING PAPER SERIES

SUNLIGHT AND PROTECTION AGAINST INFLUENZA

David Slusky
Richard J. Zeckhauser

Working Paper 24340
http://www.nber.org/papers/w24340

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2018, Revised January 2019

We thank Nils C. Wernerfelt for exceptional advice and guidance throughout this project. We are also immensely grateful to Janet Currie and Diane Alexander for their help in implementing their geographically granular influenza measure. We also thank Susan Graber, Ryan Brown, and conference participants at the 2018 American Society of Health Economists conference for their comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Sunlight and Protection Against Influenza
David Slusky and Richard J. Zeckhauser
NBER Working Paper No. 24340
February 2018, Revised January 2019
JEL No. I10,I12,I18

ABSTRACT

Recent medical literature suggests that vitamin D supplementation protects against acute respiratory tract infection. Humans exposed to sunlight produce vitamin D directly. This paper investigates how differences in sunlight, as measured over several years across states and during the same calendar month, affect influenza incidence. We find that sunlight strongly protects against influenza. This relationship is driven by sunlight in late summer and early fall, when there are sufficient quantities of both sunlight and influenza activity. A 10% increase in relative sunlight decreases the influenza index in September or October by 0.8 points on a 10-point scale. A second, complementary study employs a separate data set to study flu incidence in New York State counties. The results are strongly in accord. Remarkably, the national results are driven almost entirely by the severe H1N1 epidemic in fall 2009. That year the flu epidemic was intense, and it began early, so that September-October sunlight could play a major protective role. We also compare sunlight protection to protection produced by vitamin D supplementation in randomized trials. The sunlight effect was far greater. A plausible explanation is that exposure to sunlight is far broader, and sufficient to provide herd immunity.

David Slusky
Department of Economics
University of Kansas
335 Snow Hall
1460 Jayhawk Boulevard
Lawrence, KS 66045
david.slusky@ku.edu

Richard J. Zeckhauser
John F. Kennedy School of Government
Harvard University
79 John F. Kennedy Street
Cambridge, MA 02138
and NBER
richard_zeckhauser@harvard.edu

I. Introduction

Seasonal influenza has been with humans throughout history (Viboud and Epstein 2016). It imposes extreme costs on contemporary societies, with 2017-18 being a notable high outlier (CDC 2018). Beyond the significant discomfort to those it strikes, it saps productivity when individuals cannot work (Duarte et al. 2017) and absorbs health care resources (Schanzer and Schwartz 2013). Influenza also has less known long-range consequences. Notably, individuals exposed to influenza in utero have lower earnings as adults, are more likely to depend on government assistance (Almond 2006; Schwandt 2017), and are more likely to suffer from serious health problems later in life (Lin and Liu 2014). They are also more likely to have a heart attack (Kwong et al. 2018). Finally, influenza severity can create capacity constraints on hospitals, magnifying existing disparities in whom a hospital chooses to admit (Alexander and Currie 2017).

Influenza is a type of viral respiratory infection. Traditional public health measures to combat it include vaccination (Maurer 2009; and White 2018) and paid sick leave to keep contagious workers at home (Barmby and Larguem 2009; and Pichler and Ziebarth 2017). Coincidental reductions in interpersonal contact (such as from holiday school closings and public transportation strikes) can also reduce prevalence (Adda 2016). Finally, a recent meta-analysis shows that ingested vitamin D pills help to protect against these types of infections (Martineau et al., 2017)).[1]

This paper analyzes the potential of another mechanism for securing vitamin D: direct bodily production of vitamin D when exposed to sunlight (Holick 2007). This paper tests this mechanism's performance directly by studying population-level vitamin D production by

[1] The Martineau meta-analysis imposed stringent criteria for including a trial. This minimizes concerns about a variety of confounding factors, such as selection effects.

humans experiencing sunlight exposure. While we can ingest vitamin D from many sources, such as fish and fortified milk, passive exposure to sunlight is a much more effective source.[2] Sunlight as a source has two added benefits. First, unlike ingested vitamin D, which can become toxic at a certain concentration, the self-production mechanism does not generate toxic quantities (Holick 2007). Second, a far greater percent of the population in an area is exposed to sunlight than secures significant vitamin D for ingestion. Thus, the level of exposure is more likely to be in the range where herd immunity is significant.

Normally, flu season is in the winter, when the average sunlight level is low, and so there is not enough statistical power to identify our result. In 2009, though, the H1N1 flu epidemic hit. It peaked in the late summer and early fall. This provides us with sufficient concurrent variation that year in both sunlight and flu to study the relationship. This relationship between sunlight and flu has been studied in the broader medical literature (as by Charland et al. 2009; Grant and Giovannucci 2009; and Soebitantyo et al. 2015). Our study time period also overlaps the H1N1 outbreak in 2009. The relationship between H1N1 and Vitamin D has been studied specifically (e.g., Bruce et al. 2010; Momplaisir et al. 2012; Khare et al. 2012; Urashima et al. 2014), albeit with inconclusive results. This paper is the first to estimate the relationship by calendar month and to find that the effect is largest in late summer and early fall, when there is both substantial sunlight and sufficient influenza activity.[3] We also are the first to perform our analysis at two levels of aggregation (across states in the U.S. and across counties in New York). We find consistent results.

This paper is also the first to provide a comparison with vitamin D supplementation.

---

[2] The minimum amount of sunlight exposure (on head, neck, arm, and hands, without sunscreen) necessary to produce an effective allotment varies greatly by latitude, weather, time of year, and skin tone. In the summer it can be as short as a few minutes, whereas in the winter it can be over an hour. See http://nadir.nilu.no/~olaeng/fastrt/VitD-ez_quartMEDandMED_v2.html to calculate the minimum effective exposure time given a certain set of conditions.

[3] More broadly, a recent randomized control trial of Vitamin D supplementation found that compared to a placebo, it did not lower the incidence of cardiovascular events or invasive cancer (Manson et al. 2018).

Sunlight-created vitamin D, as opposed to ingested-supplement vitamin D, automatically tends to enhance levels broadly within a community.

II. Data

For influenza data, we used the CDC's flu index. The CDC index aggregates data reports from the individual state health department influenza surveillance points, and then harmonizes the aggregate to a consistent 10-point scale. Each point on the index represents an additional standard deviation above the mean for the ratio of visits to outpatient healthcare providers by those with symptoms of influenza, relative to all outpatient visits (regardless of symptoms). Weekly state-level data are available, from October 2008 to the present.[4] Some states, however, are missing individual weeks of data. Dropping the jurisdictions with missing flu data or sunlight data leaves us with 28 states for our primary analysis sample (CDC 2017a).[5]

We combined this flu data with the National Solar Radiation Database (NSRDB)'s daily sunlight data for 2003-2016, which covers the District of Columbia and all states but Alaska. This data represents the solar radiation for a particular set of coordinates (in watts per square meter). We calculate our primary independent variable by downloading the hourly sunlight data for the population-weighted county centroid, averaging across each month, and then constructing a county population-weighted average across counties for each state-month (Census 2010). The dataset also includes data on temperature and humidity. While our influenza data only begins in 2008, earlier data was used solely for placebo tests (NREL 2018). For one of our robustness

[4] See Appendix A for more details about the how the index is calculated.

[5] Those 28 states are: Alabama, Arizona, California, Georgia, Hawaii, Illinois, Indiana, Kansas, Maine, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Nebraska, Nevada, New Hampshire, New Jersey, Ohio Pennsylvania, Rhode Island, South Carolina, Tennessee, Texas, Vermont, West Virginia, Wisconsin, and Wyoming. As shown in Appendix Table 1, when we include all 49 states with sunlight and flu data (plus the District of Columbia), using whatever data is available for each month, we find consistent results.

checks, we also include precipitation data (which is not in NSRDB) from NOAA’s Global Surface Summary of the Day (NOAA 2017), which utilizes data from 1,218 weather stations spread throughout the United States. We assigned this data to states by matching the station closest to the population-weighted centroid for each county and then averaging with in each state-month across counties as described above (Census 2010).[6]

III. Methodology

As described above, Martineau et al. (2017)’s meta-analysis of randomized controls demonstrated significant benefits of vitamin D supplements for reducing the likelihood that an individual will contract an acute upper respiratory infection. Randomized controlled trials have served as the gold standard for epidemiological investigation. This approach follows an alternate path to methodological soundness. As an econometric study, it employs quasi-experimental variation to effectively create equivalent randomization. Implicitly, this approach controls for a wide number of variables. Moreover, it avoids the inevitable selection problems that arise when individuals must volunteer for randomized controlled trials. The current study thus employs an independent variable over which individuals had effectively no control: the deviation of a state's sunlight from its normal level.

Ideally, an econometric study would run a two-stage instrumental variable analysis, where the first stage used sunlight to predict vitamin D levels and the second stage used predicted vitamin D levels to predict influenza. Unfortunately, we lack any large scale, geo-tagged data on vitamin D levels. In its stead, our analysis employs a “reduced form” estimate of sunlight’s impact on influenza. Given that sunlight levels in a geographic area for a particular month vary randomly over the years, this provides us with a robust estimate.

[6] The correlation between the state-month average temperature variables from the two data sets is 0.9957, suggestions that there is no issue with combining weather variables from both.

Vitamin D is fat-soluble (unlike vitamin C, for example, which is water-soluble) and, therefore, has a half-life of between two weeks and two months (Mawer, Schaefer, Lumb, and Stanbury 1971; and Jones 2008). Thus, we are most interested in, and therefore calculate, the sunlight received over the month of the influenza report and the prior month. Our variable is a weighted average (by county population). Such weighting is important, because the more populous areas have a greater impact on the flu index, which is a function of the count of outpatient visits. We also calculate the monthly average flu index in each state from the weekly CDC data to get a monthly outcome variable.

We estimate the impact of the percent of deviation of sunlight (the change in log points) from its mean on deviations of the flu index from its mean as follows:

$$Flu_{smy} = \alpha + \boldsymbol{\gamma} \ln(sunlight_{smy}) + \mathbf{statemonth}_{sm} + \mathbf{year}_{y} + \varepsilon_{smy}.$$

Flu_{smy} is the flu index for state *s* in month *m* in year *y*. $Sunlight_{smy}$ refers to the average sunlight for month *m* and the prior month (as described above) for state *s* in year *y*.[7] γ is our coefficient of interest. Our preferred specification includes interaction terms (**statemonth**) for state-month fixed effects (for example, October in Kansas) and year fixed effects (for example, 2009).[8] Robust standard errors are clustered at the state level. Year fixed effects are also particularly appropriate given that the specific strains of influenza differ from year to year and vary significantly in their intensities (hence visits to the hospital if infected) and degrees of contagion. This specification follows our prior work examining the link between sunlight and vitamin D in

[7] Wernerfelt, Slusky, and Zeckhauser (2017) use data from the American Time Use Survey to show that relatively increases in sunlight increase relative time spent outdoors. We rely on their validation of this measure of sunlight and do not repeat their analysis.

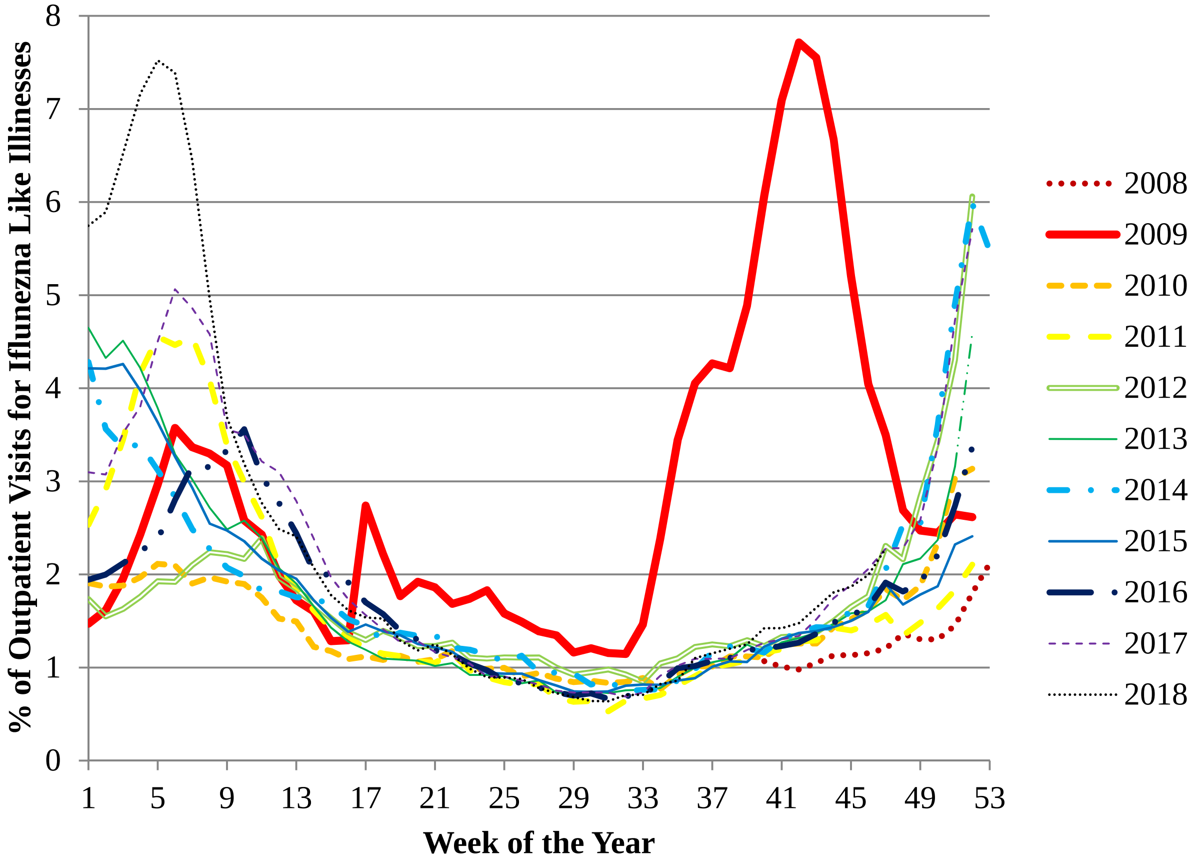
[8] Previous literature on the relationship between sunlight and flu (including Charland et al. 2009; Grant and Giovannucci 2009; and Soebitantyo et al. 2015) does not make use of fixed effects models. Given the substantial variation in latitude, weather sunlight and flu severity across states, fixed effects are crucial to ensure that estimates measure the impact of relative sunlight variation on relative flu variation, as opposed to merely identifying simple correlations.

relation to asthma. There, we found a strong protective impact of a pregnant woman's exposure to sunlight on later-in-life asthma in her child (Wernerfelt, Slusky, and Zeckhauser 2017). We also check that our results retain significance after adding a variety of weather controls, calculated analogously as county population-weighted averages, which others have found to have a significant impact on health in general and influenza in particular (including Barreca 2012; Barreca and Shimshack 2012; Deschenes 2013; Barreca, Deschenes, and Guldi 2018; Barreca et al. 2016; and Huetal, Miller, and Molitor 2017) Finally, we repeat our analysis at the county level within New York State, following the methodology of Alexander and Currie (2017) for constructing a local measure of influenza intensity.

IV. Results

As mentioned above, our time period overlaps the H1N1 epidemic of 2009. This is crucial, because of both the timing and severity of that season (CDC 2017b), as shown in Figure 1. Warmer colors are earlier years, and cooler colors later years.

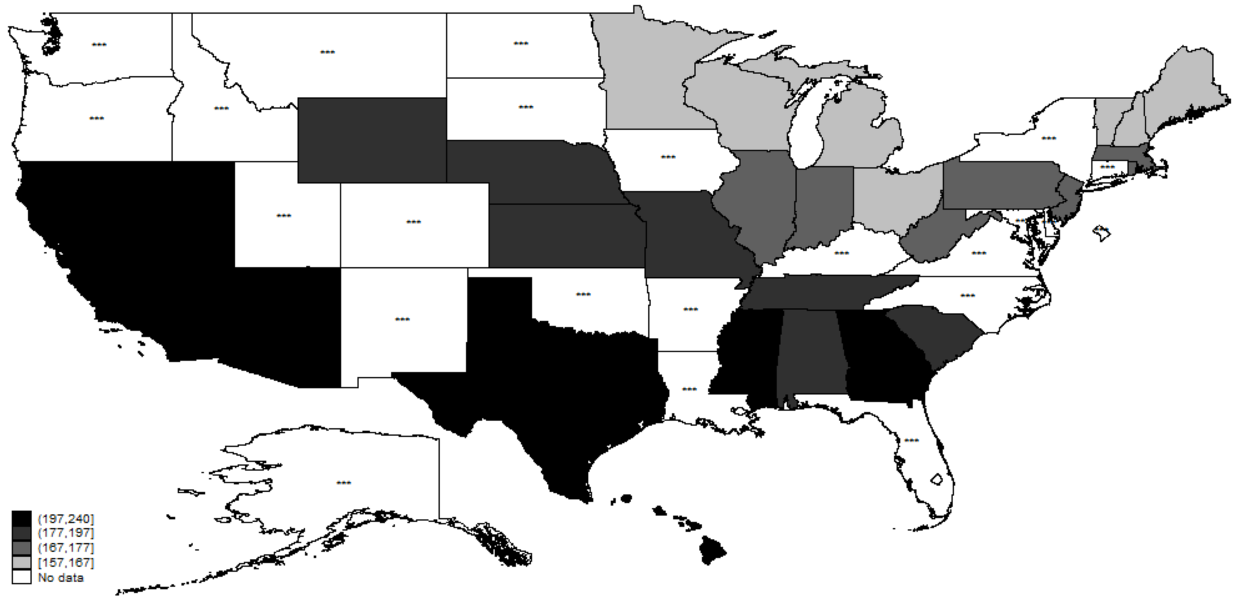
Figure 1: Weekly National Influenza Intensity



Here, we see that the 2009 influenza season was both the most severe and occurred the earliest in the year (weeks 33 to 48, corresponding to August to November). This level of greatest severity occurred during a time of the year with more sunlight overall (and therefore more room for sunlight variation) provides us with sufficient statistical power to identify our results. No other flu year was an upside outlier.

Next, we examine variation in population-weighted sunlight averages (in kilojoules per square meter per day). Figure 2 shows the three-year (2009-2016) average.

Figure 2: Population-Weighted Geographic Sunlight Variation

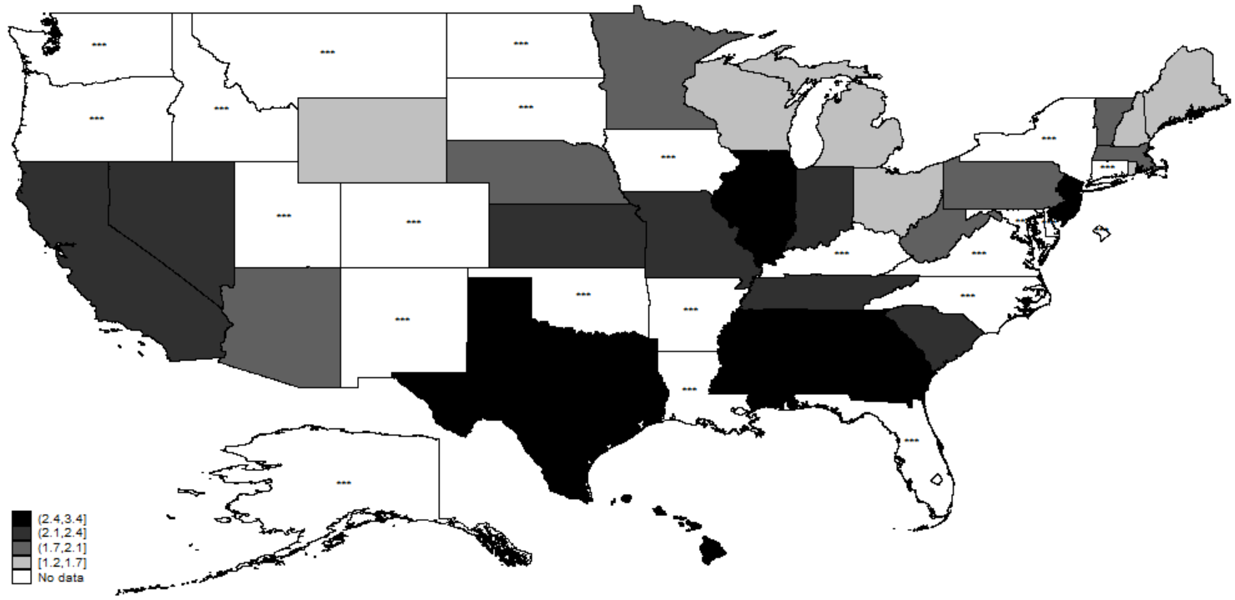


Notes: 3-year average (2009-2016) of daily county sunlight, weighted by county population. "No data" and "***" refer to incomplete influenza data for that state.

We see the expected pattern, which is that the United States is sunnier in the south and west..

Figure 3 then shows the variation by state in the average influenza index.

Figure 3: Geographic Flu Variation



Notes: 3-year average (2009-2016) of weekly state-level flu index. “No data” and “***” refer to incomplete influenza data for that state.

Here we see a very different pattern than in Figure 2. Some sunny states have high flu levels (such as Texas and California), and some low flu levels (for instance, Arizona). Moreover, some less sunny states also have high flu levels (such as Illinois), and some have low flu levels (such as Maine and New Hampshire). This suggests that other state-specific factors strongly influence influenza levels, which makes controlling for state-specific fixed effects important.

Table 1 shows summary statistics for the flu index and population-weighted average sunlight levels, as well as other weather variables (used as additional controls.)

Table 1: Summary Statistics

	(1) N	(2) Mean	(3) StDev	(4) Min	(5) Max
Flu index	2,772	2.059	2.109	1	10
Sunlight (W/m^2)	2,772	181.9	74.40	34.69	367.3
Temperature (°F)	2,772	53.56	17.83	3.841	90.02
Days/month temp <15°F	2,772	1.788	4.679	0	30.00
Specific humidity (g water vapor / kg air)	2,772	12.15	6.656	1.576	28.43
Days/month specific humidity < 6 g/kg	2,772	7.674	9.761	0	31.00
Precipitation (inches / day)	2,772	5.988	5.904	0.00118	34.65

Note: Unit of observation is a year-month for each of the 28 contiguous states that have complete flu and sunlight data.

We see that the flu index varies between 1 and 10, with an average level of 2. Sunlight also varies widely, specifically by latitude, weather, and season. Temperature and humidity also vary extensively.

Table 2 shows our initial regression results for the impact of sunlight on the influenza index, using the state-month and year fixed effect strategy described above.

Table 2: Main Results of Sunlight on Flu, All Months

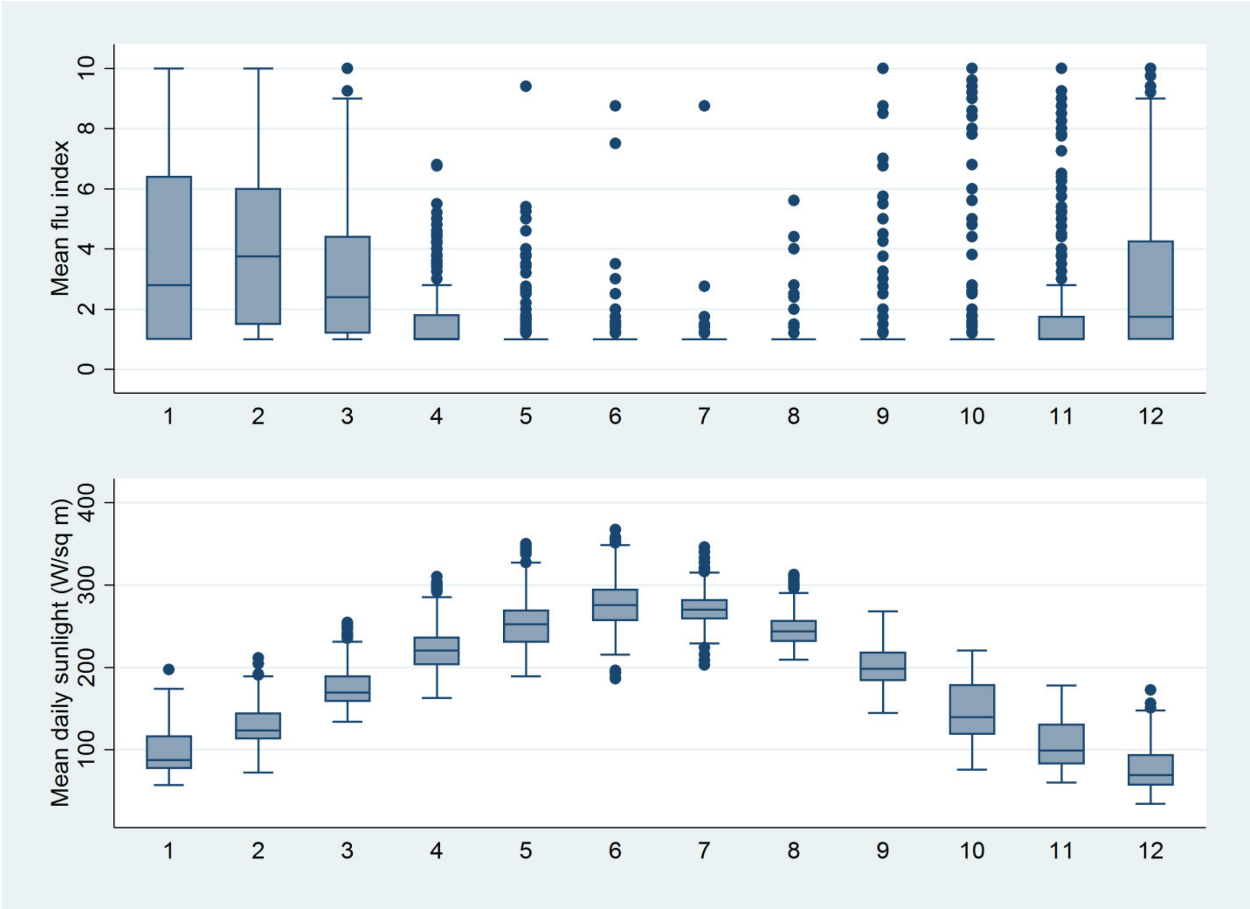
	(1)	(2)	(3)	(4)
Log sunlight for that month	-2.359***		-2.277***	
	(0.438)		(0.414)	
Log sunlight for the prior month		-0.896*	-0.550	
		(0.475)	(0.448)	
Log sunlight for that month and the prior month				-2.621***
				(0.679)
Observations	2,772	2,772	2,772	2,772
R-squared	0.095	0.085	0.095	0.091

Notes: All regressions include state-month and year fixed effects. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Column (1) of Table 2 shows that a 10% increase in relative sunlight for a month would lead to a 0.2-point decline in the influenza index for that month. In Column (2), we instead use the sunlight from the prior month (given the long half-life of vitamin D) and find a substantial effect as well. Column (3) includes each sunlight variable separately, and finds that most of the effect is on the current month. Column (4) however shows that including instead a single variable for the average sunlight over the past two months gives a coefficient that is 93% of the sum of the two coefficients in Column (3). There is uncertainty around the length of vitamin D's half-life (Mawer, Schaefer, Lumb, and Stanbury 1971; and Jones 2008). Hence, for the rest of the paper, we will include this broader two-month variable as our primary specification.

Table 2, however, includes months that have minimal influenza activity, and also months that have low levels of sunlight. Including either blunts the magnitude of the coefficients, and obscures any seasonality in the results. Figure 4 addresses this issue. It plots the ranges of influenza and sunlight by month. The top half of the figure shows that there is flu activity in the late summer, fall, and winter, but that activity is minimal in the spring and summer (except in outlier situations). The lower half shows the expected seasonal variation in sunlight levels, with large amounts of sunlight in the spring and summer and substantially less in the fall and winter.

Figure 4: Box Plots of Average Flu and Sunlight by Month



Notes: Covers the28 contiguous states that have full flu and sunlight data. Outliers are shown in blue dots.

Motivated by these plots, Table 3 re-estimates our model for each month of flu data after including the impact of that month and the prior month's sunlight. It includes only state fixed effects. Given that each column includes data for only one calendar month of each year, adding month fixed effects would have no influence.

Table 3: Month by Month[9]

	(1) Jan	(2) Feb	(3) Mar	(4) Apr	(5) May	(6) Jun	(7) Jul	(8) Aug	(9) Sep	(10) Oct	(11) Nov	(12) Dec
Log sunlight for that month and the prior month	2.314*	0.756	-0.145	0.887	0.484	-1.762	1.322	0.889	-6.752***	-5.593***	-0.594	1.435
	(1.344)	(2.054)	(2.286)	(1.111)	(0.901)	(1.735)	(1.569)	(0.823)	(2.340)	(1.650)	(2.090)	(1.501)
N	224	224	224	224	224	224	224	224	224	252	252	252
R-squared	0.734	0.381	0.328	0.223	0.126	0.102	0.076	0.172	0.519	0.856	0.665	0.554

Notes: All regressions include state and year fixed effects. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[9] Because the flu data begins in October 2008, the regressions for October, November, and December have an additional year of observations for each of the 28 states included in the primary analytic sample.

Table 3 shows that our results are being driven by September influenza (that is, August and September sunlight), and to a lesser extent by October influenza (that is, September and October sunlight). These months meet the dual requirements (as shown in Figure 4) of non-trivial level of influenza activity and still-substantial levels of sunlight. For these two months, a 10% increase in relative sunlight levels leads to a 0.6-point decline in the influenza index.

Given that Table 3 shows that the statistically significant results are found primarily in the late summer and early fall, and that Figure 1 shows that the majority of flu cases in this time of year were in the H1N1 epidemic of 2009, one might wonder whether our results are present in only 2009. Table 4 shows several analyses for that year alone, and for all of the other years.

Table 4: The Role of 2009

Panel A: Only 2009

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable	Flu Index			Difference in Flu Index from the State-Month Mean		
Log sunlight for that month and the prior month	-0.984*** (0.174)	-8.457*** (1.059)	-5.475*** (1.381)			
Difference in log sunlight for that month and the prior month from the state-month mean				-16.63*** (2.829)	-24.35*** (3.792)	-24.93*** (4.011)
N	336	84	56	336	84	56
R-squared	0.023	0.313	0.122	0.176	0.283	0.407
Months	All	Aug-Oct	Sep-Oct	All	Aug-Oct	Sep-Oct

Notes: Robust standard errors clustered at the state level in parentheses. Only 2009. *** p<0.01, ** p<0.05, * p<0.1

Panel B: Years Other Than 2009

Outcome variable	(1)	(2) Flu Index	(3)
Log sunlight for that month and the prior month	1.087 (0.806)	-0.0773 (0.341)	-0.174 (0.485)
Observations	2,352	588	392
R-squared	0.057	0.019	0.032
Months	All	Aug-Oct	Sep-Oct

Notes: All regressions include state-month and year fixed effects. 2010-2016. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

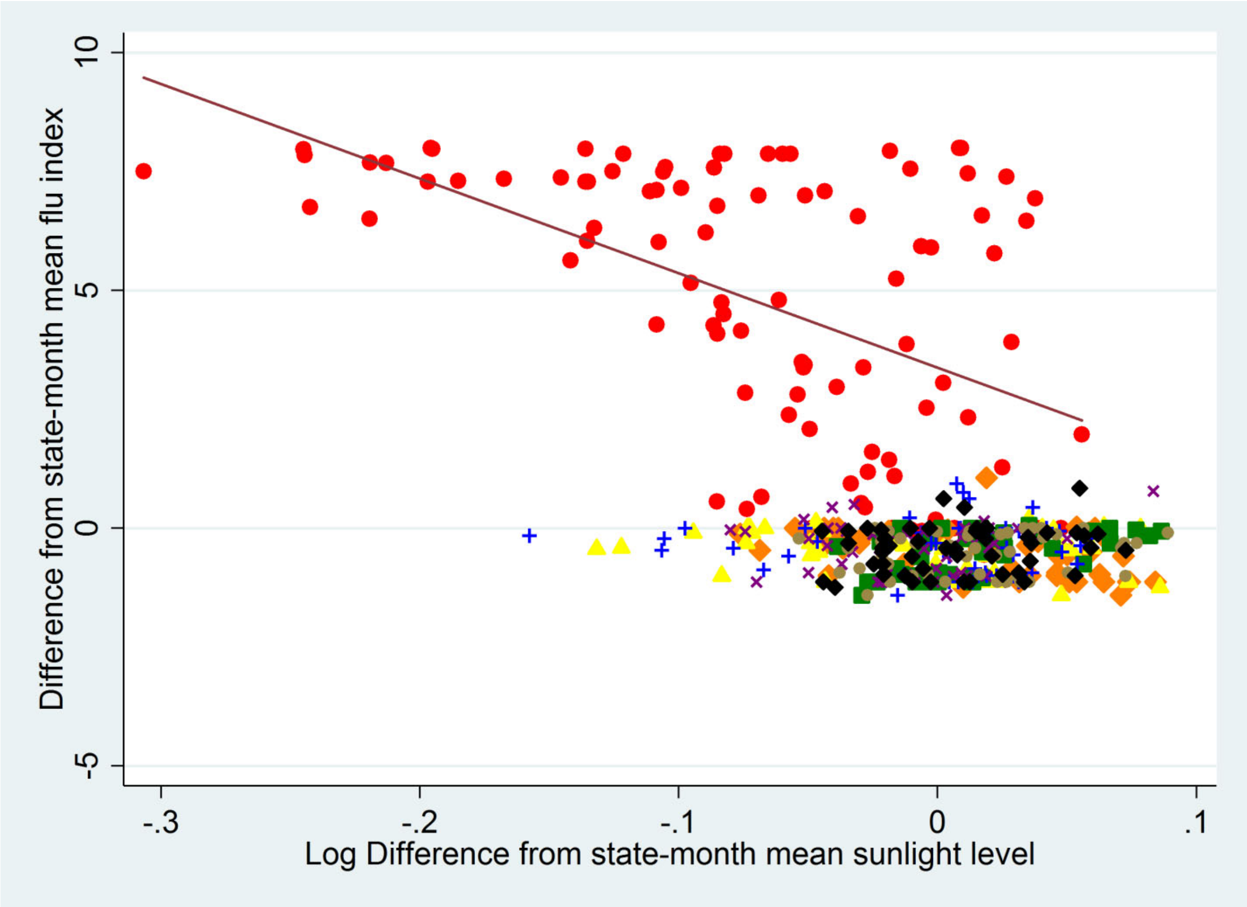
Columns (1)-(3) in Panel A show the results of regressing the flu index for a state and month on the log sunlight for that month and the prior month for only the 12 months in 2009. For any category of months (all, August-October, or Setpember-October) there is a economically and statistically significant strong negative relationship between sunlight levels and flu.

This association in theory could be endogenous, as individuals who chose to live in each of these states had prior understanding of how sunny it is each month. Therefore, following the spirit of the main analysis above, we instead calculate differences from the state-month mean value, as log points for sunlight, and by points on flu index for its incidence. This then measures not the level of sunlight but the relative difference in sunlight when compared to an average year, which should be exogenous. Here we find much larger results. The results are almost identical with and without August flu. Going forward, we therefore stick to September and October flu.

Finally, in Panel B, we re-estimate the main model in its normal form (i.e., year and state-month fixed effects), but exluding the data from 2009. Wefind no statistically significant results. This further supports our hypothesis that 2009 was a special flu year, with an early and intense season, and that it is driving our results.

We can also see this result in graphical form. Figure 5 graphs the deviations in the September and October influenza index and the log level of August/September and September/October sunlight from the mean for each state and month.

Figure 5: State-Month Deviations for Flu and Sunlight, September and October



Notes: Red Circles = 2009; Orange Diamonds = 2010, Yellow Triangles = 2011, Green Squares = 2012, Blue Pluses = 2013, Purple X's = 2014, Brown Small Circles = 2015, Black Small Diamonds = 2016. Line is linear best fit for 2009.

The horizontal axis displays our independent variable, the log of sunlight by date and month. The vertical axis graphs our dependent variable, flu index by state and month, in the difference variables in log-points calculated for Table 4 above Thus, if sunlight is protective, then the greater its level for a state and a month, the lesser will be the flu index for that state and month. As can be seen by the vertical axis in Figure 5, consistent with Figure 1 above, the 2009 flu season was substantially more severe than any of the other seasons in our sample. With this greater flu variation, a clear negative relationship emerges between relative differences in sunlight and relative differences in flu level.

To provide an additional check on the robustness of our results, Table 5 pools columns 9 and 10 (September and October) in Table 3, and adds lagged sunlight for years before the treatment period as a placebo test. If the results are robust, such lagged variables should have little or no effect. Note that here we do not include any of the weather controls that are in Table 7 below, since that would bias us against finding a statistically significant placebo result.

Table 5 shows that our primary coefficient retains its statistical significance, despite the inclusion of multiple other independent variables. Only a few coefficients on these other independent variables are significant at the 5% level. Some significance is to be expected when testing this number of hypotheses. Moreover, adding these variables hardly nudges upwards the R-square value.

Table 5: Retrospective Placebo Results for September and October Flu

		(1)	(2)	(3)	(4)	(5)	(6)
Log sunlight for that month and the prior month	Treatment year	-8.241***	-8.258***	-8.255***	-8.196***	-8.635***	-8.836***
		(1.214)	(1.236)	(1.261)	(1.267)	(1.322)	(1.400)
	Year -1		-0.283	-0.280	-0.189	-0.368	-0.926
			(0.947)	(0.978)	(0.956)	(1.003)	(1.072)
	Year -2			0.0289	0.133	-0.250	-0.596
				(0.693)	(0.721)	(0.720)	(0.840)
	Year -3				0.667	0.345	-0.143
					(0.828)	(0.884)	(0.982)
	Year -4					-1.467*	-1.932**
						(0.824)	(0.833)
	Year -5						-1.619*
							(0.906)
	Observations	476	476	476	476	476	476
	R-squared	0.707	0.707	0.707	0.707	0.708	0.710

Notes: All regressions include state-month and year fixed effects. September and October only. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Even given these findings on the powerful protective effect of sunlight, an effect supported by medical knowledge and documented with empirical analysis, there could be a concern that our results are picking up some other kind of environmental variation. One possibility would be temperature, given that it is known to have health effects (Deschenes 2013; Barreca, Deschenes, and Guldi 2018; Barreca et al. 2016; and Huetal, Miller, and Molitor 2017). Thus, following Barreca, Deschenes, and Guldi (2018) and Wernerfelt, Slusky, and Zeckhauser (2017), we now control for the number of days per month that a state experiences extreme cold (daily low temperature below 15°F). Such control is merited, because the influenza virus can survive better between hosts at lower temperatures (Polozov et al. 2008). Absolute humidity can also play a role in influenza mortality. Prior work identifies a negative nonlinear relationship between humidity and influenza, where levels below 6 g of water vapor per kg of air had a substantial impact (per Barreca 2016; Barreca and Shimshack 2012).[10] Finally, we also include precipitation, as it is possible that a lack of sunlight is acting through this channel.

The results after adding these additional controls are shown in Table 6.

[10] Specific humidity is not directly provided in the NSRDB data, so we calculated it using the available information on dew point and atmospheric pressure and the Tetens equation. See http://snowball.millersville.edu/~adecaria/ESCI241/esci241_lesson06_humidity.pdf for the necessary formulas.

Table 6: Results Controlling for Other Weather Measures, September and October

	(1)	(2)	(3)	(4)	(5)	(6)
Log sunlight for that month and the prior month	-8.382*** (1.362)	-8.350*** (1.224)	-10.17*** (1.279)	-9.160*** (1.184)	-8.383*** (1.144)	-11.74*** (1.435)
Controls (past two months):						
Log temperature	X					X
Days per month below 15°F		X				X
Log specific humidity			X			X
Days per month specific humidity is below 6 g/kg				X		X
Log precipitation					X	X
Observations	476	476	476	476	476	476
R-squared	0.707	0.713	0.714	0.713	0.707	0.724

Notes: All regressions include state-month and year fixed effects. September and October only. Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

As shown in Table 6, adding these weather controls have minimal effect on our primary finding, and the exclusion of them if anything is biasing our main results toward zero.

We now turn to replicating our results at the sub-state level. First, we need to show that our results are consistent for at least some subsets of states.[11] Table 7 shows our results stratifying by quartiles of overall average sunlight levels.[12] Here we include all states (even those

[11] As we'll show below, our results are robust to omitting each state one at a time, and so are not driven by any individual state.

[12] 1st Quartile: Delaware, District of Columbia, Maine, Michigan, Minnesota, Montana, New Hampshire, North Dakota, Oregon, Vermont, Washington, and Wisconsin

2nd Quartile: Connecticut, Idaho, Illinois, Indiana, Iowa, Massachusetts, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, South Dakota, and West Virginia,

3rd Quartile: Alabama, Arkansas, Kansas, Kentucky, Maryland, Missouri, Nebraska, North Carolina, South Carolina, Tennessee, Utah, Virginia, and Wyoming with missing weeks of influenza data) and all months, both to maximize statistical power and because the state we will eventually look within (New York) is one of these states with incomplete data.

Table 7: Results Stratified by Average Sunniness of State

	(1)	(2)	(3)	(4)
Quartile of Sunniness	1st (least sunny)	2nd	3rd	4th (sunniest)
Log sunlight for that month and the prior month	-2.537***	-4.099***	-3.997***	-0.288
	(0.651)	(0.739)	(0.877)	(1.490)
Observations	1,086	1,274	1,280	1,188
R-squared	0.125	0.123	0.110	0.134
Number of state-months	139	156	156	144
States	12	13	13	12

Notes: All regressions include state-month and year fixed effects and weather controls for that month and the prior month (log temperature, days per month below 15°F, log specific humidity, days per month specific humidity is below 6 g/kg, and log precipitation). Robust standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Except the sunniest quartile of states, our results are broadly consistent across quartiles. We can therefore proceed to analyze within New York (which falls in the second quartile).

The analysis below follows Alexander and Curie (2017), who construct a ZIP-code level weekly flu measure for New Jersey. Here, we use SPARCS hospital discharge data (New York State Department of Health 2015) for all of New York for October 2008 (the earliest month we have CDC flu data) to June 2014 (the last year for which we have discharge data and the last quarter for which we have bed data).

The method can be briefly described as follows:

4th Quartile: Arizona, California, Colorado, Florida, Georgia, Hawaii, Louisiana, Mississippi, Nevada, New Mexico, Oklahoma, and Texas

1. Take all emergency department discharges and all inpatient discharges with an emergency department indicator (since those admitted from the ED drop out of the ED file)

2. Keep those emergency discharges with a influenza flu diagnoses (CCS[13] code of 123) and inpatient discharges with an influenza diagnosis that was present on arrival.

3. Use the admitted date to assign to an epidemiological week (always Sunday-Saturday, which is the CDC standard).[14]

4. Sum for all New York State for each week and compare to the flu index:

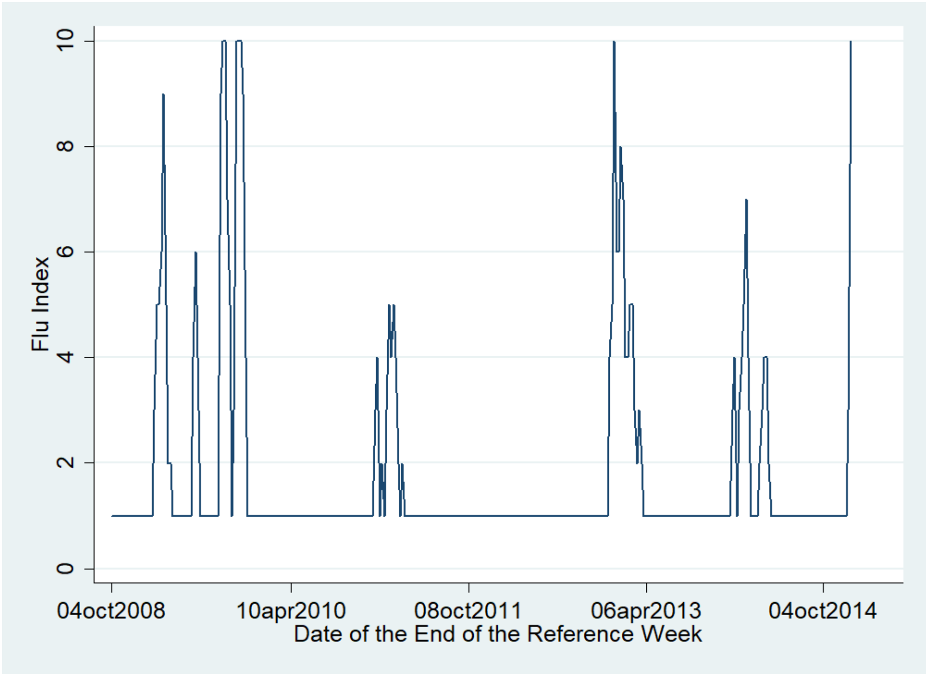
To validate this measure, following Alexander and Currie (2017) in Figure 6 we compare the weekly New York State flu index from the CDC (2017a) with the total number of influenza admissions that week in New York.

[13] https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp

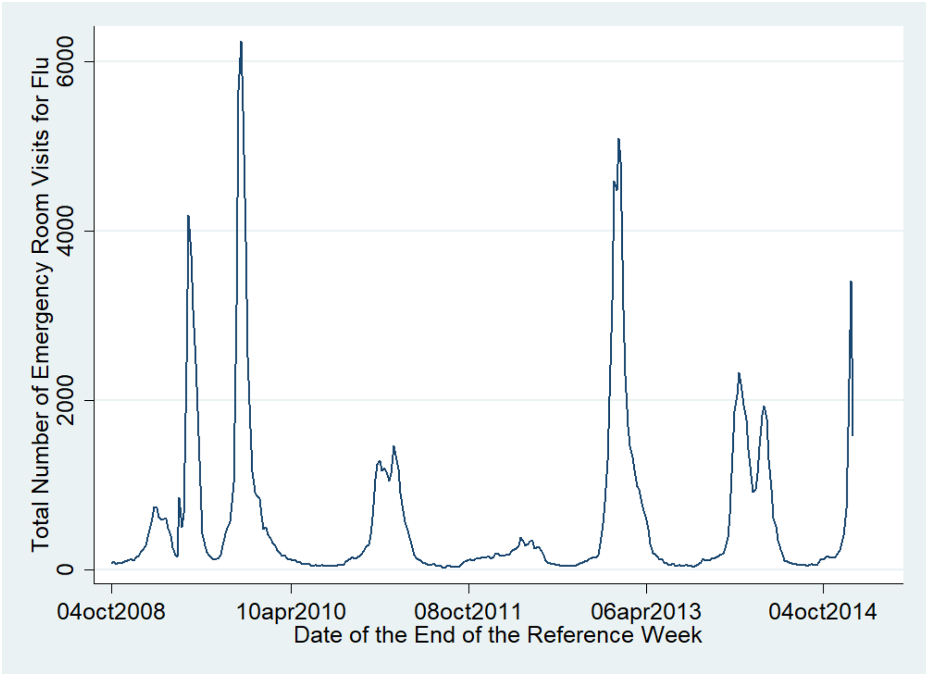
[14] https://wwwn.cdc.gov/nndss/document/MMWR_Week_overview.pdf, using the Stata command "epiweek"

Figure 6: Comparison of CDC and Discharge Flu Measures for New York State

Panel A: CDC Flu Index



Panel B: Hospital Emergency Room Visits for Flu



While our discharge-based measure is more continuous than the CDC's measure, the two track each other remarkably well. This is all the most significant given that the CDC measure is based on outpatient office visits and not emergency department visits.

We can use these counts of influenza discharges to construct a county-level measure, again following Alexander and Currie (2017). Briefly, the steps are as follows:

1. For each hospital, merge in bed data (New York State Department of Health 2016)[15] and divide the number of admissions in that week by the number of beds to get the per bed admissions rate.

2. For each county centroid (Census 2010), calculate the great circle distance to the geocoded coordinates of each hospital's address.

3. For each hospital within 100 miles, divide the per bed influenza rate by the distance between the county and the hospital and then sum to get the county level influenza flu index

4. Average the county-week level index over a month per the main analysis above

We then merged this county-level influenza measure with the county-level sunlight and weather data from NSRDB. New York has 62 counties, and so we actually have more units here than in the main analysis.

[15] The current number of beds is available on the New York State Department of Health's website. Historical information through the second quarter of 2014 was obtained in response to an email request.

Table 8: New York State County-Level Analysis for Hospital-Based Influenza Measure

	(1)	(2)	(3)	(4)	(5)	(6)
Log sunlight for that month	-0.464***	-0.454***				
	(0.0622)	(0.0624)				
Log sunlight for the prior month		-0.0953**				
		(0.0367)				
Log sunlight for that month and the prior month			-0.654***	-2.715***	-3.741***	-2.817***
			(0.0821)	(0.837)	(1.064)	(0.502)
Weather Control					X	X
Dropping Outlier County						X
Observations	4,832	4,832	4,832	744	744	732
R-squared	0.108	0.109	0.108	0.276	0.339	0.512
Months	All	All	All	Sep-Oct	Sep-Oct	Sep-Oct

Notes: 2008-2014. All regressions include state-month and year fixed effects. For Columns (5) and (6), weather controls are for that month and the prior month and include log temperature, days per month below 15°F, log specific humidity, days per month specific humidity is below 6 g/kg. Robust standard errors clustered at the county level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Across a wide variety of sunlight measures, sample months, and controls, we find very consistent results to those above at the national level.

Table 9 is analogous to Table 4, looking at only 2009 to see if our results appear in only that year, given that it is the source of our late summer / early fall flu variation, and at the rest of the years excluding 2009.

Table 9: The Role of 2009, New York Counties

Panel A: Only 2009

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable		Flu Index			Difference in Flu Index from the State-Month Mean	
Log sunlight for that month and the prior month	-0.147***	-2.358***	-2.997***			
	(0.0219)	(0.387)	(0.545)			
Difference in log sunlight for that month and the prior month from the state-month mean				-2.971***	-4.457**	-13.15***
				(0.298)	(1.980)	(3.853)
N	744	186	124	744	186	124
R-squared	0.009	0.310	0.275	0.098	0.036	0.202
Months	All	Aug-Oct	Sep-Oct	All	Aug-Oct	Sep-Oct

Notes: Robust standard errors clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Panel B: Years Other Than 2009

	(1)	(2)	(3)
Outcome variable		Flu Index	
Log sunlight for that month and the prior month	0.0156	-0.0790***	-0.130***
	(0.0391)	(0.0144)	(0.0332)
Observations	4,088	929	620
R-squared	0.068	0.141	0.174
Months	All	Aug-Oct	Sep-Oct

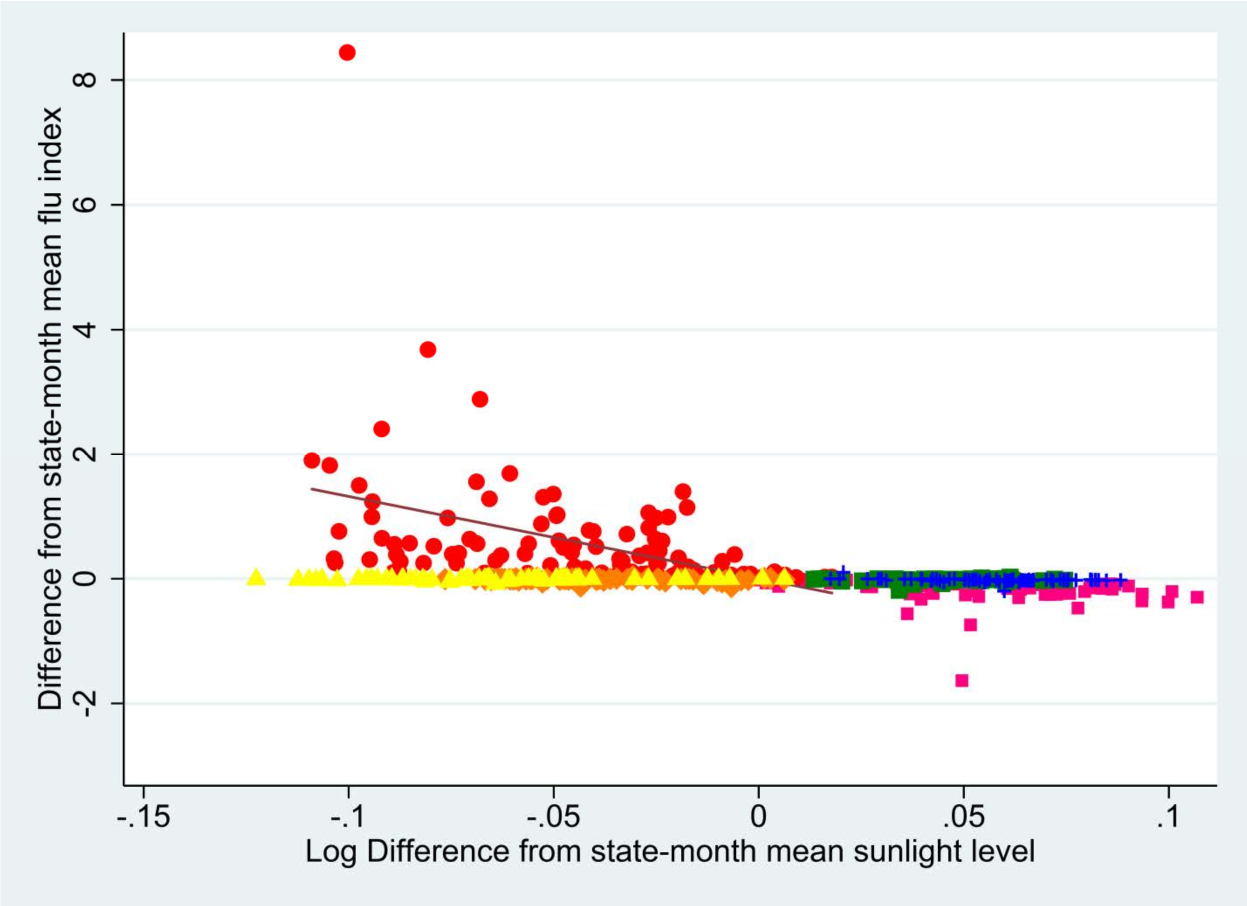
Notes: 2008, 2010-2014. All regressions include state-month and year fixed effects. Robust standard errors clustered at the county level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Here, in Panel A, we see consistent results, both in the cross section and using differences from county-month averages across years. In Panel B, in Column (1), we see no statistically significant result, analogous to Table 4 above. In Columns (2) and (3), despite the exclusion of

2009, we still find a statistically significant result that relatively greater levels of fall sunlight led to relatively lower influenza intensity. The coefficients, however, are a tiny fraction of what they were for 2009.

We can also create Figure 7, analogous to Figure 5, comparing the differences in sunlight and flu within counties across years.

Figure 7: State-Month Deviations for Flu and Sunlight, September and October, Counties



Notes: Pink Small Squares = 2008, Red Circles = 2009; Orange Diamonds = 2010, Yellow Triangles = 2011, Green Squares = 2012, Blue Pluses = 2013, Purple X's = 2014, Brown Small Circles = 2015, Black Small Diamonds = 2016. Line is linear best fit for 2009.

Here we see a similar relationship to above, where the variation is driven by 2009 (consistent with Figure 1 and Figure 5 above) and is downward sloping (more sunlight & less flu). The results are robust to dropping the one outlier at the top left of the graph (Erie County)—see Table 8, Column (6).

Additional Robustness Checks

The tables in Appendix B conduct additional robustness checks. Appendix Table B1 repeats the Table 2 analyses, but includes an unbalanced panel of all contiguous states, Hawaii, and D.C. (i.e., even those with missing influenza data in some weeks). It finds a comparable result. It also employs a linear specification and finds strongly statistically significant results, though obviously at different coefficient magnitudes.

Appendix Table B2 drops each of one of the 28 states in the primary specification, one at a time, to show that the main result is robust to the exclusion of any one particular state. Appendix Table B3 performs the analysis for only sunlight from each day of the week (e.g., the average sunlight on Sundays in a given month), and finds that any day's sunlight has an impact.

Finally, Appendix Table B4 performs the analysis as a weekly level, with the primary variable calculated as the average over the previous eight weeks, and with state-week fixed effects. The results are consistent with those found at the month level, and including sunlight from a year prior yields a statistically insignificant placebo result.

V. Discussion

Impact on welfare. We can attempt to estimate the impact of our results on welfare. As described above, each point on the influenza index represents an additional standard deviation above the mean of the non-flu week's ratio of outpatients presenting with symptoms of influenza to all outpatients (CDC 2017a). The data is also available on the actual outpatient counts, though not broken down at the state level (CDC 2017b).

As described above, the flu index indicates the number of standard deviations above the non-influenza mean of the share of outpatients who exhibit influenza symptoms. In the 2005-2008 “pre-period,” this mean share is 1.03%, and the standard deviation is 0.394 percentage points.[16]

Figure 5 shows that the range of relative sunlight levels for September and October within state-months across years is roughly plus or minus 0.05 log points, that is, 10 percentage points. Thus, our coefficient for log sunlight shown in Table 5 corresponds to a 0.8241-point reduction in the influenza index, which can be interpreted as 0.8241 standard deviations. Given that one standard deviation is 0.394 percentage points, 0.8241 standard deviations represents 0.32 percentage points.

The average annual total number of all outpatients in September and October (weeks 35 to 43) in our study years (2009-2016) (from CDC 2017b) is 6,342,726. A 0.32 percentage point reduction would produce 20,595 fewer cases.

To translate this into a dollar amount, we need two additional pieces of information. First, Molinaria et al. (2007) estimate that the total cost of seasonal influenza is $87 billion per year.[17] Second, again using the CDC (2017b) data, the average annual number of influenza patients for 2009-2016 is 697,025. Our reduction of 20,595 is 3.0%, which gives us an approximate monetary equivalent savings of $2.6 billion.

<u>Herd Immunity</u>. Giving 100 people in a town of perhaps 10,000 people a vitamin D supplement will offer extremely flu externalities of protection. But give that same town extra sunlight, and most of the community will produce vitamin D, thereby conveying an externality of

[16] See Appendix A for additional calculation details.
[17] This estimate includes the cost of hospitalization and outpatient visits, lost earnings, and life-years lost. It does not include disutility from having the flu.

protection that triggers herd protection against influenza, a highly communicable disease.[18] Positing that supplements and sunlight-produced vitamin D are equivalently powerful, that externality could massively increase the magnitude of the protective effect.

To test this conjecture, we compared our results to those for vitamin D supplementation. By contrast, the Martineau et al. (2017) meta analysis of 25 randomized controlled trials of vitamin D supplementation found an adjusted odds ratio of only 0.88 for acute respiratory tract infections. A likely contributor to the disparity relates to externalities promoting herd immunity when sunlight is the protective factor. It is possible, of course, that part or all of the disparity is because sunlight produces greater and/or more effective vitamin D than supplements.

<u>Virus Deactivation</u>. Sunlight can also protect against influenza via a path apart from the production of vitamin D. Ultraviolet light deactivates the virus directly (Sagripanti and Lytle 2007). The data in this paper provides no direct way to assess the relative contributions of these two mechanisms. However, we can be confident that the vitamin D path is consequential, as the Martineau et al. (2017) meta analysis demonstrates.

[18] This herd immunity obviously would also benefit those who do not go outdoors, as the more outdoorsy people with whom they come in contact would be less likely to be infected and contagious.

VI. Conclusion

Sunlight, likely operating through the well-established channel of producing vitamin D , plays a significant role in flu incidence. A recent meta-analysis of 25 randomized controlled trials of vitamin D supplementation (Martineau et al. 2017) demonstrated significant benefits of such supplements for reducing the likelihood that an individual will contract an acute upper respiratory infection. The current study considers sunlight as an alternate, natural path through which humans can and do secure vitamin D. This study's findings reinforce the Martineau et al. findings.

Our main finding is that incremental sunlight in the late summer and early fall has the potential to reduce the incidence of influenza. It did so dramatically in 2009, when the flu came early – giving the more powerful sunlight of the later summer the opportunity to protect – and in a year when it was particularly powerful. Apart from its methodological contributions, this study reinforces the long-held assertion that vitamin D protects against acute upper respiratory infections. One can secure vitamin D through supplements, or through a walk outdoors, particularly on a sunny day.

References

Adda, J. 2016. Economic Activity and The Spread of Viral Diseases: Evidence From High Frequency Data. Quarterly Journal of Economics, 131(2): 891-941.

Alexander, D. and J. Currie. 2017. Are publicly insured children less likely to be admitted to hospital than the privately insured (and does it matter)? Economics and Human Biology 25: 33-51.

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

Anderson, R.M., and R.M. May. 1990. Immunisation and Herd Immunity. The Lancet 335(8690): 641-645.

Barmby, T. and M. Larguem. 2009. Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. Journal of Health Economics 28: 1012-1017.

Barreca, A.I. 2012. Climate change, humidity, and mortality in the United States. Journal of Environmental Economics and Management 63(1):19–34.

Barreca, A.I., K. Clay, O. Deschenes, M. Greenstone, and J.S. Shapiro. 2016. Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the 20th Century. Journal of Political Economy 124(1): 105-159.

Barreca, A.I., O. Deschenes, and M. Guldi. 2018. Maybe Next Month? The Dynamic Effects of Ambient Temperature on Fertility. Demography, 55(4):1269-1293.

Barreca, A.I. and J.P. Shimshack. 2012. Absolute Humidity, Temperature, and Influenza Mortality: 30 Years of County-Level Evidence from the United States. American Journal of Epidemiology 176(7): S114-S122.

Bartlett, M.S. 1957. Measles Periodicity and Community Size. Journal of the Royal Statistical Society 120(1): 48-70.

Black, Francis L. 1996. Measles endemicity in insular populations: Critical community size and its evolutionary implication. Journal of Theoretical Biology 11(2): 207-211.

Bruce, D, J..H Ooi, S. Yu, and M.T. Cantorna. 2010. Vitamin D and host resistance to infection? Putting the cart in front of the horse. Experimental Biology and Medicine 235(8): 921–927.

CDC. 2017a. Influenza-Like Illness (ILI) Activity Level Indicator Determined by Data Reported to ILINet https://gis.cdc.gov/grasp/fluview/main.html

CDC 2017b. National, Regional, and State Level Outpatient Illness and Viral Surveillance. https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html

CDC 2017c. Overview of Influenza Surveillance in the United States. https://www.cdc.gov/flu/weekly/overview.htm

CDC 2018. Transcript for CDC Update on Flu Activity. https://www.cdc.gov/media/releases/2018/t0126-flu-update-activity.html

Census 2010. US Population Weighted Center. https://www2.census.gov/geo/docs/reference/cenpop2010/county/CenPop2010_Mean_CO.txt

Charland, K.M.L., D.L. Buckeridge, J.L. Sturtevant, F. Melton, By. Reis, K.D. Mandl, and J.S. Brownstein. 2009. Effect of environmental factors on the spatio-temporal patterns of influenza spread. Epidemiol. Infect 137:1377–1387.

Deschenes, O. 2013. Temperature, human health, and adaptation: A review of the empirical literature. Energy Economics 46(C): 609-619.

Duarte, F., S. Kadiyala, S.H. Masters, and D.Powell. 2017. The Effect of the 2009 Influenza Pandemic on Absence from Work. Health Economics, 26(12):1682-1695.

Grant, W.B., and E. Giovannucci. 2009. The possible roles of solar ultraviolet-B radiation and vitamin D in reducing case-fatality rates from the 1918–1919 influenza pandemic in the United States. Dermato-Endocrinology, 1(4):215-219.

Holick, M.F. 2007. Vitamin D deficiency. New England Journal of Medicine, 357(3): 266-281.

Huetal, G., N.H. Miller, and D. Molitor. 2017. Adaptation and the Mortality Effects of Temperature Across U.S. Climate Regions. NBER Working Paper No. w23271.

Jones, G. 2008. Pharmacokinetics of vitamin D toxicity. Am J Clin Nutr, 88(2):582S-586S.

Khare, Drirh, M Godbole, Nachiket, Pawar, Shailesh, Mohan, Vishwa, Pandey, Gaurav, Gupta, Sushil, Tripathi, Deepak, Dhole, Tapan, Godbole, Madan. 2012. Calcitriol [1, 25[OH]2 D3] pre- and post-treatment suppresses inflammatory response to influenza A (H1N1) infection in human lung A549 epithelial cells. European journal of nutrition. 52(4): 1405-1415.

Kwong, J.C., K.L. Schwartz, M.A. Campitelli, H. Chung, N.S. Crowcroft, T. Karnauchow, K. Katz, D.T. Ko, A.J. McGeer, D. McNally, D.C. Richardson, L.C. Rosella, A. Simor, M. Smieja, G. Zahariadis, and J.B. Gubbay. 2018. Acute Myocardial Infarction after Laboratory-Confirmed Influenza Infection. New England Journal of Medicine, 378(4): 345-353.

Lin, M.J. and E.M. Liu. 2014. Does in utero exposure to Illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. Journal of Health Economics, 37: 152-163.

Manson, J.E., et al. 2018. Vitamin D Supplements and Prevention of Cancer and Cardiovascular Disease. Forthcoming at New England Journal of Medicine.

Martineau, A.R., D.A. Jolliffe, R.L. Hooper, L. Greenberg, J.F. Aloia, P. Bergman, G. Dubnov-Raz, S. Esposito, D. Ganmaa, A.A. Ginde, E.C. Goodall, C.C. Grant, C.J. Griffiths, W. Janssens, I. Laaksi, S. Manaseki-Holland, D. Mauger, D.R. Murdoch, R. Neale, J.R. Rees, S. Simpson, I. Stelmach, G.T. Kumar, M. Urashima, C.A. Camargo. 2017. Vitamin D supplementation to prevent acute respiratory tract infections: systematic review and meta-analysis of individual participant data. Forthcoming at BMJ.

Momplaisir, F., Frank, I., Meyer, W., Kim, D., Kappes, R., & Tebas, P. 2012. Vitamin D Levels, Natural H1N1 Infection and Response to H1N1 Vaccine among HIV-Infected Individuals. Journal of AIDS & Clinical Research, 3(4), 152.

Maurer, J. 2009. Who has a clue to preventing the flu? Unravelling supply and demand effects on the take-up of influenza vaccinations. Journal of Health Economics, 28: 704-717.

Mawer, E.B., K. Schaefer, G.A. Lumb, and S.W. Stanbury. 1971. The metabolism of isotopically labelled vitamin D3 in man: the influence of the state of vitamin D nutrition. Clin Sci, 40: 39-53.

Molinaria, NA.M., I.R. Ortega-Sanchez, M.L. Messonnier, W.W. Thompson, P.M. Wortley, E. Weintraub, C.B.Bridges. 2007. The annual impact of seasonal influenza in the US: Measuring disease burden and costs, Vaccine 25(27):5086-5096.

NREL. 2018. National Solar Radiation Database. https://nsrdb.nrel.gov/

NOAA. 2017. Federal Climate Complex Global Surface Summary of Day Data. ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt

New York State Department of Health. 2015. Statewide Planning and Research Cooperative System (SPARCS) https://www.health.ny.gov/statistics/sparcs/

New York State Department of Health. 2016. NY State Health Profiles. https://profiles.health.ny.gov/hospital/alpha

Pichler, S., and N.R. Ziebarth. 2016. The Pros and Cons of Sick Pay Schemes: Testing for Contagious Presenteeism and Noncontagious Absenteeism Behavior. Journal of Public Economics, 156: 14-33.

Polozov, I. V., L. Bezrukov, K. Gawrisch and J. Zimmerberg. 2008. Progressive ordering with decreasing temperature of the phospholipids of influenza virus. Nature Chemical Biology 4(4): 248-255.

Sagripanti, J.L. and C.D. Lytle. 2007. Inactivation of Influenza Virus by Solar Radiation. Photochemistry and Photobiology 83: 1278–1282.

Schanzer, D.L., and Schwartz, B., 2013. Impact of Seasonal and Pandemic Influenza on Emergency Department Visits, 2003–2010, Ontario, Canada. Acad Emerg Med 20: 388–397.

Schwandt, H. 2017. The Lasting Legacy of Seasonal Influenza: In-utero Exposure and Labor Market Outcomes. IZA Working Paper 10589.

Soebitantyo, R.P., D. Gross, P. Jorgensen, S. Buda, M. Bromberg, Z. Kaufman, K. Prosenc7, M. Socan, T.V. Alonso, M.A. Widdowson, R.K. Kiang. 2015. Associations between Meteorological Parameters and Influenza Activity in Berlin (Germany), Ljubljana (Slovenia), Castile and León (Spain) and Israeli Districts. PLOSOne. 10(8):e0134701.

Urashima M, Mezawa H, Noya M, Camargo CA Jr. 2014. Effects of vitamin D supplements on influenza A illness during the 2009 H1N1 pandemic: a randomized controlled trial. Food Funct. 5(9):2365-70.

Viboud, C., and S.L. Epstein. 2016. First flu is forever. Science 354 (6313): 706-707.

Wernerfelt, N., D.J.G. Slusky, and R.J. Zeckhauser. 2017. Second Trimester Sunlight and Asthma: Evidence from Two Independent Studies. American Journal of Health Economics, 3(2): 227-253.

White, C. 2018. Measuring the Social and Externality Benefits of Influenza Vaccination. *Mimeo*, available at http://coreywhite.net/s/White_Flu_24Aug2018.pdf.

ZIP Code Database. 2013 https://www.unitedstateszipcodes.org/zip-code-database/

Appendix A

We use the weekly count of outpatient visits (both total and only those due to influenza) from the CDC (2017b) along with the documentation in the ILI data (CDC 2017c) to perform calculations regarding the influenza index. That index corresponds to the number of standard deviations the share of outpatient visits that report influenza symptoms that week differs when compared to all non-influenza weeks. A "non-influenza week" is defined as a week during which that week and its preceding week had fewer than 2% of all outpatient visits to healthcare providers indicating influenza.

As our study period is 2008-2016, we use the October 2005-September 2008 period as a "pre-period" to calibrate our index. We begin with the formal start of the season, which the CDC defines as week 40 (the first week of October). Unfortunately, whereas the ILI data (CDC 2017a) is available at the state level, the outpatient visit count data is only available nationally. Therefore, we perform our calculations at that level.

Nationally, of the 156 weeks in October 2005-September 2008, 108 fit the above definition of "non-influenza." The mean share for those 108 weeks is 1.03%, and their standard deviation is 0.39 percentage points.

Given this, the method for calculating the influenza index is now to take all weeks, calculate the z-score[s] (that is, number of standard deviations above or below the mean), and then apply the following index definition:

Flu index =

1	if	$Z < 0$
int(Z) +2	if	$0 < Z < 8$
10	if	$Z > 8$

So, in the interior range of the index, we can consider an additional index point as an additional standard deviation.

Appendix B

Appendix Table B1: Results with States with Some Months Missing Flu Values, All Months

	(1)	(2)	(3)	(4)
States	All	All	Non Missing	Non Missing
Months	All	Sept & Oct	All	Sept & Oct
Log sunlight for that month and the prior month	-2.792***	-5.766***		
	(0.418)	(0.948)		
Sunlight for that month and the prior month			-0.00998**	-0.0407***
			(0.00374)	(0.00677)
Observations	4,835	826	2,772	476
R-squared	0.103	0.735	0.087	0.700

Notes: All regressions include state-month and year fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table B2: Dropping One State at a Time

State dropped	Log sunlight for that month and the prior month		R-squared
Alabama	-7.843***	(1.31)	0.707
Arizona	-8.508***	(1.349)	0.719
California	-8.689***	(1.257)	0.719
Georgia	-8.204***	(1.404)	0.712
Hawaii	-8.138***	(1.307)	0.746
Illinois	-8.153***	(1.31)	0.715
Indiana	-8.190***	(1.327)	0.713
Kansas	-8.040***	(1.318)	0.71
Maine	-7.714***	(1.19)	0.735
Massachusetts	-8.119***	(1.301)	0.728
Michigan	-8.141***	(1.291)	0.727
Minnesota	-8.201***	(1.332)	0.724
Mississippi	-8.134***	(1.363)	0.71
Missouri	-8.031***	(1.331)	0.711
Nebraksa	-8.058***	(1.317)	0.718
Nevada	-8.737***	(1.254)	0.716
New Hampshire	-7.998***	(1.291)	0.731
New Jersey	-8.516***	(1.282)	0.727
Ohio	-8.407***	(1.321)	0.727
Pennsylvania	-8.484***	(1.347)	0.716
Rhode Island	-8.216***	(1.315)	0.722
South Carolina	-8.364***	(1.355)	0.712
Tenneesee	-8.161***	(1.368)	0.708
Texas	-7.972***	(1.321)	0.71
Vermont	-8.362***	(1.331)	0.73
West Virginia	-8.537***	(1.314)	0.724
Wisconsin	-8.461***	(1.312)	0.712
Wyoming	-8.389***	(1.358)	0.726

Notes: N=459. September and October only. All regressions include state-month and year fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table B3: By Days of the Week

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log sunlight for that month and the prior month, only for:							
Sunday	-1.689**						
	(0.653)						
Monday		-3.186***					
		(0.606)					
Tuesday			-3.978***				
			(0.670)				
Wednesday				-1.336**			
				(0.621)			
Thursday					-3.139***		
					(0.607)		
Friday						-2.921***	
						(0.606)	
Saturday							-2.604***
							(0.542)
Observations	476	476	476	476	476	476	476
R-squared	0.669	0.683	0.694	0.667	0.686	0.681	0.677

Notes: September and October only. All regressions include state-month and year fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table B4: Weekly Level Analysis

Months	(1) All	(2) Sept & Oct	(3) All	(4) Sept & Oct	(5) All	(6) Sept & Oct
Log sunlight for that month:	-2.475*** (0.543)	-7.292*** (0.886)				
Log sunlight for that month and the prior month			-1.573** (0.660)	-7.413*** (1.126)	-1.583** (0.675)	-7.404*** (1.150)
Log sunlight for that month and the prior month, one year earlier					-0.140 (0.594)	-1.327 (1.032)
Observations	12,068	2,072	12,068	2,072	12,068	2,072
R-squared	0.078	0.667	0.072	0.645	0.072	0.646

Notes: All regressions include state-week and year fixed effects. *** p<0.01, ** p<0.05, * p<0.1