NBER WORKING PAPER SERIES

SECOND TRIMESTER SUNLIGHT AND ASTHMA: EVIDENCE FROM TWO INDEPENDENT STUDIES

Nils Wernerfelt
David Slusky
Richard Zeckhauser

Working Paper 20599
http://www.nber.org/papers/w20599

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2014

We thank Amitabh Chandra and Scott Weiss, MD, for detailed comments and suggestions, and seminar participants at the Leonard Davis Institute, Princeton's Center for Health and Wellbeing, and MIT. We also thank Bob Krasowski of the Centers for Disease Control, and Jean Roth and the rest of the staff at NBER for outstanding assistance compiling our datasets, and Diane Alexander, Asaf Bernstein, Hoyt Bleakley, Anne Case, Alma Cohen, Janet Currie, David Cutler, Moshe Hazan, Fernanda Marquez Padilla, Sarah Moshary, Paul Murphy, Stephen Murphy, Joseph Newhouse, Jessica Pickett, Analia Schlosser, Benjamin Sommers, Heidi Williams, and especially Adam Sacarny for valuable comments. Wernerfelt gratefully acknowledges the support of an NSF Graduate Research Fellowship. This paper explored allergy and asthma emergency rates in children using outpatient data from Arizona and New Jersey, State Emergency Department Databases (SEDD), Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the Research Data Center, the National Center for Health Statistics, the Centers for Disease Control and Prevention, or the National Bureau of Economic Research.

Second Trimester Sunlight and Asthma: Evidence from Two Independent Studies
Nils Wernerfelt, David Slusky, and Richard Zeckhauser
NBER Working Paper No. 20599
October 2014
JEL No. I12,I18,J13

ABSTRACT

One in twelve Americans suffers from asthma and its annual costs are estimated to exceed $50 billion. Simultaneously, the root causes of the disease remain unknown. A recent hypothesis speculates that maternal vitamin D levels during pregnancy affect the probability the fetus later develops asthma. In two large-scale studies, we test this hypothesis using a natural experiment afforded by historical variation in sunlight, a major source of vitamin D. Specifically, holding the birth location and month fixed, we see how exogenous within-location variation in sunlight across birth years affects the probability of asthma onset. We show that this measurement of sunlight correlates with actual exposure, and consistent with pre-existing results from the fetal development literature, we find substantial and highly significant evidence in both datasets that increased sunlight during the second trimester lowers the subsequent probability of asthma. Our results suggest policies designed to augment vitamin D levels in pregnant women, the large majority of whom are vitamin D insufficient, could be very cost-effective.

Nils Wernerfelt
MIT Department of Economics
77 Massachusetts Avenue, E19-750
Cambridge, MA 02139
nilsw@mit.edu

David Slusky
Department of Economics
Princeton University
and
Department of Economics
The University of Kansas
1460 Jayhawk Boulevard
Lawrence, KS 66045
david.slusky@ku.edu

Richard 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

The worldwide incidence of asthma imposes staggering physical and financial costs. Asthma afflicts more than 300 million individuals worldwide and kills more than 250,000 people annually (Cruz et al., 2007). The resulting cost - measuring both direct treatment expenditures and indirect productivity loss - is enormous, estimated at about $56 billion annually in the U.S. alone (Barnett and Nurmagambetov, 2011). Simultaneously, asthma's root causes appear varied and remain poorly understood. A recent hypothesis with growing support speculates that maternal vitamin D levels during pregnancy may affect the probability that the unborn child later develops asthma (Weiss and Litonjua, 2011). If validated, this hypothesis presents a low-cost way to reduce asthma incidence, resulting in both improved quality of life for millions of individuals and reductions in net healthcare spending.

To assess this hypothesis, we examine the impact of local weather conditions while *in utero* on the probability of developing asthma later in life. Specifically, our approach holds fixed the location and month of birth, and then analyzes how exogenous within-location variation in sunlight levels across birth years affects the probability of asthma onset. Though we cannot observe maternal vitamin D levels directly, the motivation for our approach stems from the well documented connection between vitamin D levels and sunlight. Namely, Americans obtain over 90% of their vitamin D from sunlight exposure (Holick, 2004), and we provide evidence in our Appendix that the sunlight variation we are exploiting correlates with actual exposure, but not with other factors that might be affecting asthma incidence.

In two independent datasets, we find highly significant evidence that an increase in sunlight in an individual's location during the second trimester *in utero* reduces his/her probability of becoming asthmatic. This result is consistent with pre-existing research from the fetal development literature, which has found substantial evidence suggesting vitamin D is particularly important during this period for asthma pathogenesis.

We report the results from our two datasets separately. Our first dataset (Study I) draws on individual level data on the state, month, and year of birth for more than 260,000 Americans. We find that doubling the amount of sunshine[1] during the second trimester lowers the probability of an asthma diagnosis by 1.15 percentage points, which at a mean incidence of 11.49% represents a 10.0% reduction. In our second dataset (Study II), we look for an effect of sunlight on a separate, more cost-relevant margin - hospital discharges from asthma - on the county level using data on more than 2.1 million births aggregated into about 3,000 birth month-county cohorts. Here we find that doubling the amount of sunlight during the second trimester reduces the per capita rate of asthma emergencies by 2.21 percentage points, which is a 21.3% reduction at the mean.[2] Both sets of results pass several robustness checks detailed in the Appendix (Section 6).

Current estimates are that nearly three out of every four adults are either vitamin D deficient or insufficient (Ginde, Lui, and Camargo, 2009).[3] Among pregnant women, multiple smaller scale studies have reported rates that are if anything slightly higher (Bodnar et al., 2007; Holmes et al., 2009; Johnson et al., 2011). Our results suggest these numbers may be more costly than previously known and that policies designed to boost

[1] e.g., about the difference between the average summer in Nevada and that in New Hampshire.
[2] This estimate is significantly different from the previous one, but this is understandable given the datasets look at two different margins. We report the results here in terms of doubling sunlight for sake of comparison, but we describe each result in more detail in Section IV.
[3] Deficiency is defined as less than 10 ng/ml blood; insufficiency is 10-30 ng/ml.

maternal vitamin D levels during the second trimester -- either through increased usage of (or adherence to) vitamin D supplements or increased sun exposure -- may greatly enhance welfare.[4]

The rest of the paper is organized as follows. Section II provides the background from both the medical and economics literatures. Section III describes our two datasets in greater detail. Section IV presents our results and Section V concludes.

II. Background

Our paper builds off of two separate literatures, and so we provide the relevant background for the two sequentially before discussing the connection between sunlight and vitamin D in more detail.

II. A. Prior Medical Literature: Many studies in both animals and humans point to a role for maternal vitamin D levels during pregnancy in asthma pathogenesis.[5,6] One stream of this literature has tried to find evidence of a clinical effect. Multiple studies on human subjects in different parts of the world have found associations between low maternal vitamin D levels during pregnancy and asthma incidence in the child (Camargo et al., 2007; Devereux et al., 2007; Erkolla et al., 2009; Miyake et al., 2010; Belderbos et al.,

[4]A year's worth of vitamin D supplements typically costs less than $20 online. In addition, exposure to as little as ten minutes of direct sunlight a day -- far less time than it takes to sunburn -- is adequate for most individuals to obtain their daily recommended dosage (Holick, 2007). In contrast, the average asthma-related hospitalization lasts three days and can cost between $12,000 and $24,000 (Milet et al., 2007).
[5] There have also been studies recently that look at treating current asthmatics with vitamin D (e.g., Castro et al., 2014). This is a related, but distinct hypothesis from what we consider in this paper.
[6] We note at the outset that the evidence is sufficiently strong that two randomized-controlled trials have been started to look at maternal vitamin D supplementation and during pregnancy and asthma outcomes. These studies are the Maternal Vitamin D Supplementation to Prevent Childhood Asthma (VDAART) and Vitamin D Supplementation During Pregnancy for Prevention of Asthma in Childhood (ABCvitaminD); their ClinicalTrials.gov identifiers are NCT00920621 and NCT00856947, respectively.

2011). By design though, these studies can only report associations, and other associational studies report either no correlation or a positive one between low maternal vitamin D levels and later asthma incidence (e.g., Gale et al., 2007). Separately, in a controlled study on animals, rats born to mothers who were experimentally deprived of vitamin D exhibit several lung abnormalities similar to those found in human asthmatics (Gaultier et al., 1984).

A second stream of literature has been more developmentally focused and tries to determine when in a human pregnancy maternal vitamin D levels are most important. Human lung development occurs in five stages: embryonic (roughly 4-7 weeks post-conception), pseudoglandular (7-17 weeks), canalicular (17-26 weeks), saccular (27-36 weeks), and alveolar (36 weeks to about 2 years old) (Burri, 1984; Weiss and Litonjua, 2011). Several studies find suggestive evidence that maternal vitamin D levels could be important during part of the pseudoglandular and canalicular stages.[7,8] Not only is vitamin D present in fetal lung tissue during much of this period (Brun et al., 1987), but it is also known to affect either the growth or gene expression of several types of cells that are active in the lungs at this time (Lunghi et al., 1995; Stio et al., 1997; Bossé, Maghni, and Hudson, 2007). At week 11 of the pregnancy, many vitamin D related genes are suddenly activated that are known to be involved in lung development and differentially expressed in asthmatics (Kho et al., 2013). These studies illuminate a pathway through which vitamin D deficiency could lead to the abnormal airway development in asthma

[7] It is worth emphasizing that the fetal supply of vitamin D is completely determined by the mother's (e.g., Heaney et al., 2003), and so statements about vitamin D in fetal lungs during this period reflect the mother's supply.

[8] We note that other studies have also reported results consistent with the third trimester (e.g., Glasgow and Thomas, 1977; Phokela et al., 2005). We do not find evidence in support of this, but note that there may well be an effect during that period that is not large enough to observe in our data.

(Weiss and Litonjua, 2011). Based on these results, we look for evidence within weeks 11-26 post conception, which overlaps relatively closely with the second trimester.

II. B. Prior Economic Literature: Our paper relates most closely to the literature within economics on the long-run impacts of events early in life. Much of that research shows that outcomes of considerable interest -- such as educational achievement, income, and a host of health conditions -- are strongly influenced by events early in life (e.g., Heckman, 2006; Bleakley, 2007; Fletcher, Green, and Neidell, 2010; Bharadwaj, Løken, and Neilson, 2013) or *in utero* (e.g., Black, Devereux, and Salvanes, 2007; Almond, Edlund, and Palme, 2009). Almond and Currie (2011) show how this empirical evidence of the "fetal origins hypothesis" can fit within a standard economics modeling framework. Namely, using a constant elasticity of substitution production function with health investments at different ages as inputs, they show that complementarities between health shocks and later investments could amplify small initial differences across individuals.[9]

Our paper also relates to several prior economic analyses that rely on weather as a proxy variable for a variable of interest. For example, Maccini and Yang (2009) report on several positive long-term outcomes for women who were exposed to more rainfall early in life, where the presumed causal channel operates through increased agricultural production and hence income. Jensen (2000) uses rainfall as a proxy for agricultural conditions to study investments in children, and Oster (2004) uses temperature as a proxy for economic conditions to study the prevalence of witchcraft trials. Similar to these

[9] This is an adaption of a model from Heckman (2007).

studies, we do not observe maternal vitamin D levels directly,[10] but run our analyses using sunlight instead (which we provide evidence is correlated). Clearly, direct measurement would be ideal, but our approach still enables us to provide evidence of the hypothesized effect.

III.C. Sunlight and Vitamin D: As mentioned, humans obtain the vast majority of their vitamin D supply from sunlight (in fact, a nickname for vitamin D is "the sunshine vitamin"). Though there are many dietary sources of vitamin D - such as several fish species or a range of fortified dairy products - it would take about 30 glasses of vitamin D fortified milk to produce the same amount of vitamin D as from an average of 5-10 minutes of direct sunlight exposure (Holick, 2007).

Similarly, in the northern hemisphere, vitamin D levels are known to follow strong seasonal cycles of being highest in the summer and lowest in the winter as the number of sunlight hours varies (and the pattern is reversed in the southern hemisphere; Chen, 1999). Isolated data from small scale studies have similarly found seasonal fluctuations in the vitamin D level of both pregnant women and their newborns (Disanto et al., 2013; Bodnar et al., 2007). Finally, on a separate but related note, rickets (which is caused by vitamin D deficiency) was documented more than a century ago to follow similar seasonal cycles (Kassowitz, 1897). For identification purposes, the variation in sunlight we exploit is within-location rather than seasonal, but we include this section to simply convey a sense of the tight connection between sunlight and vitamin D.

[10] There are a few datasets that contain information on vitamin D levels, the largest of which is the NHANES, but even that does not have enough geographical and seasonal variation for our purposes.

III. Data

This paper presents two separate studies. Table I provides an overview of each. The text then elaborates on the description of each dataset.

Table I. Overview of our two studies.

	Study I	Study II
Unit of observation	Individual	Cohort
Level of observation	Each individual's state, month, and year of birth	Each cohort's county, birth month, and year of birth
Weather data	Hours of sunlight, at the state, month, and year level	Average intensity of sunlight at the county, month, and year level
Outcome variable	The individual's self-reported asthma condition	Per-capita number of emergency department discharges for asthma for that cohort
Years	1914-1987	1999-2009
N	264,701 individuals	2.1 million people grouped into 3,036 birth month-county-year cohorts

We employ two analyses first to see if our main effect replicates, and second since each study has distinct advantages. Study I has the advantage of having individual level data, with sunlight aggregated by state, over a national sample. Study II, by contrast, operates at the more granular county level for sunlight exposure, uses more up-to-date data in a narrower time frame, and has a more definitive indication of asthma from emergency department discharges. We note that the two studies have no overlap in data (with respect to either time frame or data sources). This creates a highly independent replication.

Study I: Individual level data on the state and month of birth (as well as health status and controls) for more than 260,000 Americans were obtained from the 1997-2008 National Health Interview Survey (NHIS). This data was merged with historical weather data from the National Oceanic and Atmospheric Administration (NOAA).

An adult from the NHIS data was coded as asthmatic if he/she responded affirmatively to the question "Have you EVER been told by a doctor or other health professional that you had asthma?" and similarly for a child below 18 years old if a knowledgeable adult in the household answered affirmatively to an equivalent question about the child. Negative responses were coded as a 0; all other responses were omitted (e.g., refused to answer, not ascertained, don't know).[11] The publicly available NHIS data also provides information on each individual's year and month of birth, as well as demographic data such as gender and ethnicity. The restricted-use data we used included each individual's state of birth. Thus, for every individual in our sample we knew his/her state, month, and year of birth.

Historical sunlight data were collected from 240 NOAA weather stations located around the country. The data range from 1891-1987, though not all stations cover the entire time period.[12] Each station measured the total number of hours of sunlight at its location each month. For example, we can see in our data that Asheville, North Carolina had 188 hours of sunlight in January of 1977. For states with more than one weather station reporting in any month, the average within-state correlation in sunlight hours was quite high (0.86), and so monthly data were averaged in such cases. In Section 2 of the

[11] The NHIS is described as the "principal source of national asthma prevalence data for the United States" (National Center for Environmental Health, 2014). It still relies on self-reported asthma status though, and so our second study measures asthma using hospital discharge data.

[12] The average duration of a weather station was just above 60 years, and the vast majority were reporting over continuous stretches.

Appendix we provide figures to demonstrate the variation in sunlight at the level of identification (Figures A.1-A.2).

Study II: Our second dataset looks at birth month-county-year cohorts, not individual data. Sunlight data came from the National Solar Radiation Database (NSRD) at the U.S. Department of Energy, asthma hospital discharge data from the Health Care Utilization Project, and birth records from the CDC's Vital Statistics.

The sunlight data shows for each 10 km x 10 km grid cell a measure of sunlight intensity at hourly intervals (though for our analysis we use the monthly average) for 1998-2009.[13] To map the NSRD to counties, since there is minimal sunlight variation within a county for a particular month and year, we matched the coordinates of the population weighted county centroid to the 10 km x 10 km gridded cell that contains it. As for Study I, Section 2 of the Appendix contains visual depictions of the variation in sunlight used in this study at the level of identification (Figures A.3-A.4).

We use hospital discharge data from the State Emergency Department Database, which records visits to the emergency room that did not result in an admission. We were able to obtain data from two states: New Jersey and Arizona. The discharge data for these states contains the patient's county of residence, birth month, and birth year, information critical for our analysis. We limit our sample to children aged 0-10 years old when

[13] For each 10 km x 10 km grid cell, Geostationary Operation Environmental Satellites measure solar radiation reflected back from any cloud cover. Ground radiation can then be calculated from subtracting reflected radiation from total radiation, presumably using latitude and day of the year to calculate total radiation (Perez et al. 2002). These ground radiation estimates are then integrated over an hour, since the most granular data available is at the hour-day-month-year level. Any gaps in the satellite data are corrected using a long fill method with data from other years for that time and location, smoothing endpoints. The units for each hour are watt-hours per square meter, or equivalently an average instantaneous rate of watts per square meter.

discharged, and we assume that their county of residence is also the county of birth.[14] Discharges are coded as "asthma" if the diagnosis code they have falls into that category as defined by the clinical classifications software.

As mentioned, this study aggregates data to the county-birth month-birth year cohort. We thus normalize the discharge data by the number of live births in each county in the respective birth month and birth year, using vital statistics data from the CDC.[15]

Whereas Study I focuses on incidence of asthma (without any measure of exacerbations), Study II fundamentally covers exacerbations of asthma, without any real measure of incidence across the population. A high per capita asthma emergency discharge rate could be caused by more individuals in that cohort having asthma and therefore more individuals having emergencies, or it could be caused by the same types of individuals who would have had asthma in a different cohort having more severe asthma and therefore more emergencies. In this way, the two studies complement each other, and show that the relationship between sunlight *in utero* and asthma is robust to the choice of outcome measure.

Given this background, Table II provides means and general summary statistics for our two datasets.[16]

[14] We limit the sample to ages 0-10 to minimize the effects of cohort differences across time (e.g., changes in maternal smoking habits). With regards to our county of birth assumption, we note that even if there were significant movement, it would enter our regressions as classical measurement error as the movement should not be correlated with whether a county was having a more or less sunny month relative to its mean. The measurement error would not be classical only if individuals moving somewhere somehow caused that place to be relatively sunnier than it had been before.

[15] Unfortunately, the publicly available data covers only counties with populations greater than 100,000. For birth years 1999-2009, data is publicly available for 18 of New Jersey's 21 counties, but only 5 of Arizona's 15 counties. This should not bias the results of this analysis, since there is not likely to be a relationship between county population and within-county-month annual sunlight variation. Furthermore, there is a concern that less populated counties will have a count of emergency asthma or allergy emergency discharges censored at 0, which will bias a linear regression model. Omitting those counties helps avoid that bias.

Table II. Summary statistics for Studies I and II.

Study I		Study II	
N	264,701	*N*	3,036
		N (total births)	2,119,857
Percent asthmatic:	11.49%	*Per capita rate of asthma emergencies*	0.104[17]
Average year of birth:	1956.06	*Average year of birth:*	2004
Fraction female:	56.0%	*Average cohort share female:*	48.8%
Major Ethnicities:		*Major Ethnicities (average cohort share)*	
White	81.2%	White	79.1%
African American	15.7%	African American	13.0%
Other	2.8%	Other	n/a
Hours of Sunlight:		*Sunlight intensity (W/m^2):*	
Mean ± Std. Dev.		*Mean ± Std. Dev.*	
1st Trimester:	673.74±196.79	1st Trimester:	182.88±65.99
2nd Trimester:	676.58±197.28	2nd Trimester:	182.42±65.49
3rd Trimester:	668.82±197.24	3rd Trimester:	182.10±65.55

Construction of the sunlight variable: Neither dataset enables us to trace the weather exactly back to the day of birth. Given this coarseness in our data, we cannot determine the precise weeks of the pregnancy to which a sunlight measure pertains. However, we

[16] For Study I, our full sample consists of 595,093 people, but analyzing the means of our variables of interest across the entire dataset and those for whom we have asthma data reveals no major differences (Table A.1 in the Appendix, Section 1).

[17] The per capita rate of asthma emergencies is calculated by the count of 2005-2009 emergency department discharges with an asthma diagnosis for a particular birth month - birth year - county cohort, divided by the number of individuals born in that cohort. This rate is larger in magnitude than one might expect since multiple discharges for the same individual are counted separately and since it pools multiple years. Data limitations make it impossible to track the same individual across discharges and years, which would be necessary to calculate an individualized rate.

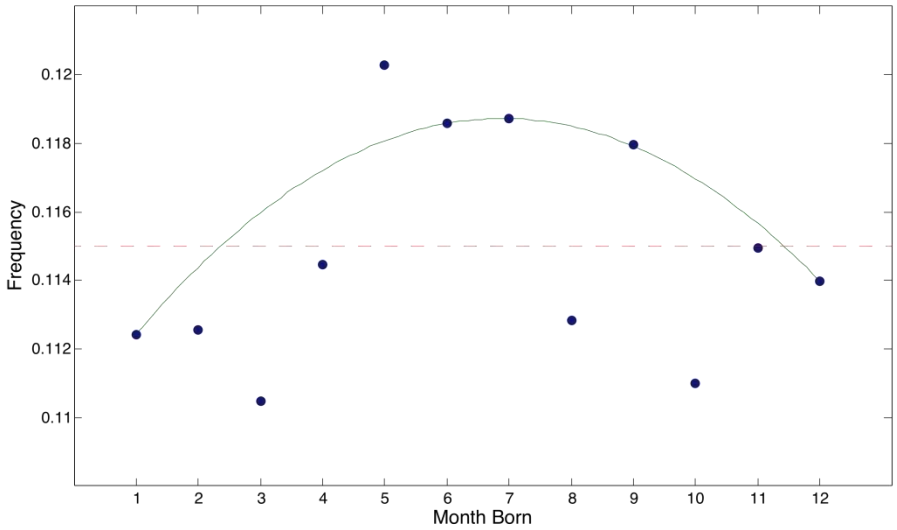
can give a range. For example, in the Appendix, Section 3 we run regressions at the month level (i.e. instead of trimesters on the right hand side we have months). Consider if we found that the month two months prior to birth was significant. For individuals born on the first day of their month of birth this would be the second to last month of the pregnancy, but for those born on the last day of the month, it would be the third to last. We thus bound our results accordingly and in that case would say that we find evidence of significance within the last two to three months of the pregnancy.

For both studies, our trimester measurements for sunlight sum three adjacent months. Thus, our third trimester starts with the sunlight from one month prior to the birth month and adds to it the amounts in the two preceding months. We define the other trimesters correspondingly. We start one month prior to the birth month so that our second trimester measurement aligns better with the late pseudoglandular and canalicular stages.

IV. Methodology and Results

At a very broad level, we begin by comparing the distributions of normal and asthmatic births across the calendar year utilizing no controls, employing the dataset from Study I. The distributions differ significantly (Pearson's Chi-squared test, $\chi^2=25.73$, df=11, $p=0.0071$), with asthmatic births more common roughly during the summer. Figure I shows the frequency of asthmatic births by month (e.g., among January births, 11.24% in our sample were later diagnosed with asthma).

Figure I. Probability of being diagnosed with asthma as a function of birth month.



Note: The dashed red line denotes the sample average, the blue dots represent the data, and the green line fits the data with a cubic spline (knots at months 1, 6 and 12).

Given the seasonal variation in weather across most of the nation, these data suggest a potential connection between sunlight and asthma. However, other factors could contribute to or produce this pattern. For example, it is known that seasonality of birth correlates with race (e.g., Bound and Jaeger, 2001; Buckles and Hungerman, 2013; Currie and Schwandt, 2013). Therefore, to the extent that race is correlated with asthma incidence, racial differences in the season of birth could conceivably produce this pattern.

To study our research question in greater detail, we now turn to our main results, which use the aforementioned natural experiment to rigorously test the sunlight hypothesis. Our approach fixes the birth location and month, and then assesses how within-location sunlight variation across years affects the probability of asthma.[18] We

[18] For this approach to provide meaningful evidence on a possible vitamin D connection, we need to assume that this variation in sunlight correlates with vitamin D levels. Though this may seem obvious, to provide supportive evidence, we merged our data from Study II with that from the American Time Use

report results as linear probability models with our binary measure of asthma as the dependent variable and measures of sunshine and our controls as independent variables.

Study I: Given the individual-level data, this study's regressions take the following form:

$$ASTHMA_{ismy} = \alpha + \lambda_{sm} + \rho_y + \gamma \ln(SUN_{smy}) + X_i'\beta + \varepsilon_{ismy},$$

where $ASTHMA_{ismy}$ is the asthma status of individual i born in state s in month m and year y. λ_{sm} are state of birth interacted with month of birth fixed effects; ρ_y are year of birth fixed effects; $\ln(SUN_{smy})$ is the log-transformed measurement of sunlight; and the X_i are individual controls for race and gender.

In words, this specification proceeds as follows. Individuals born in July in Georgia in 1978 received a certain exposure to sunlight *in utero*; individuals born in July in Georgia in 1979 received a different exposure. We assess how that variation in exposure affects the probability of developing asthma, taking into account a possible national trend in asthma diagnoses across years with the ρ_y.

Our main results are displayed below in Table III. We report our analysis using log-transformed sunlight to express our results in terms of percent changes, but an analysis using linear sunlight yields very similar outcomes.[19]

Survey, to show that the variation we exploit correlates with individuals' time spent outside. Since we view this as supportive and not our main results, we refer the reader to the Appendix, Section 4, for more on this point.

[19] See Table A.6.

Table III. Regression output of main results from Study I.

Dependent variable:	(1) Asthma	(2) Asthma	(3) Asthma
Log Hours of Sun 3rd Trimester	0.00123	0.00156	0.00140
	(0.00711)	(0.00707)	(0.00711)
Log Hours of Sun 2nd Trimester	-0.01615***	-0.01592***	-0.01653***
	(0.00595)	(0.00590)	(0.00596)
Log Hours of Sun 1st Trimester	0.00274	0.00318	0.00360
	(0.00726)	(0.00728)	(0.00717)
State of Birth*Month of Birth Fixed Effects	x	x	x
Year of Birth Fixed Effects	x	x	x
Race controls		x	x
Gender			x
N	264,533	264,533	264,533
R^2	0.0081	0.0089	0.01073

Notes: The results indicate that greater sunlight exposure during the second trimester reduces the probability of being diagnosed with asthma later in life. For the coefficient on the second trimester sunlight, p=0.0091, 0.0095, and 0.0079 for Columns (1)-(3), respectively. All regressions are done with robust standard errors clustered at the state level. (*** denotes significance at 0.01)

This Table's striking finding is that sunlight exposure during the second trimester has a large and significant effect on the probability that a child is later diagnosed with asthma. More specifically, doubling the amount of sunshine during this period lowers the subsequent probability of asthma by 1.15 percentage points.[20] Thus, at the mean incidence rate of 11.49%, such an increase in sunlight would lower the probability of asthma by around one tenth. In contrast, neither sunlight during the first or third trimester is significant. To put these results in context, the average number of hours of sunlight in the summer in Nevada is about twice that for New Hampshire. Similarly, a baby born in

[20] This number comes from multiplying the coefficient on the second trimester in Column (3) by ln(2): -0.01653*ln(2)=-0.0115.

December in New Mexico has on average about three times as much sunlight during the second trimester as a baby born in May in Michigan.

Bounding the time window as mentioned in the Data section, these results provide evidence for significance inside weeks 10-26 of the pregnancy. This is a close match with the window of weeks 11-26 we specified earlier.[21]

Our results pass various robustness checks detailed in the Appendix, Section 6 (Table A.8). Namely, if we add an additional pregnancy's worth of sunlight data both before conception and after birth to our main regressions, none of these extra terms is significant while our main effect persists. Similarly, repeating our analysis using an indicator variable for low birth weight or whether an individual's father graduated from high school (neither of which is hypothesized to be connected to *in utero* vitamin D) as the dependent variable, none of the sunlight coefficients is significant.

Study II: The main regression used in this study is:

$$ASTHMA_{cmy} = \alpha + \lambda_{cm} + \rho_y + \gamma \ln(SUN_{cmy}) + \varepsilon_{cmy},$$

where $ASTHMA_{cmy}$ is the per-capita number of emergency department discharges for asthma in county c for individuals born in month m and year y. λ_{cm} are county interacted with month of birth fixed effects; ρ_y are year of birth fixed effects; and $\ln(SUN_{cmy})$ is the log-transformed measurement of sunlight.

In words, what this specification does is as follows. The cohort of individuals born in Mercer County, New Jersey in March, 2006 received a certain exposure to sunlight *in utero*; individuals born there in the March, 2007 cohort received a different

[21] Given our data from both studies are at the month level, we repeat the analysis at this finer level in the Appendix, Section 3, to see if we can learn more about when during this period sunlight exposure is most important (Table A.2).

exposure. We are determining how that variation affects the incidence rate of emergency department discharges for asthma for those cohorts, taking into account possible trends across birth years that would affect all counties. Our main results from this study are in Table IV. As above, we report our analysis using log-transformed sunlight, but rerunning with linear sunlight yields very similar outcomes.[22]

Table IV. Regression output of main results from Study II.

Dependent variable:	(1) Asthma	(2) Asthma
Log Average Sunlight intensity 3rd Trimester	0.0114 (0.0244)	-0.0226 (0.0339)
Log Average Sunlight intensity 2nd Trimester	-0.0319*** (0.0098)	-0.0598*** (0.0209)
Log Average Sunlight intensity 1st Trimester	0.0205 (0.0197)	-0.0124 (0.0197)
County of Birth*Month of Birth Fixed Effects	x	x
Year of Birth Fixed Effects	x	x
Birth years	1999-2009	1999-2004
N	3,036	1,656
R^2	0.7624	0.7467

Notes: We run the same regression on the full sample and a subset of the birth years.[23] The results indicate that greater sunlight exposure during the second trimester reduces that cohort's per capita rate of emergency department discharges for asthma later in life. For the coefficient on the second trimester sunlight, p=0.004, and 0.009 for Columns (1)-(2), respectively. All regressions are done with robust standard errors are clustered at the county level. (*** denotes significance at 0.01)

[22] See Table A.7.

[23] One might be concerned that there are fewer years of discharge data available with 2005-2009 for cohorts born after 2004 since they were alive for less of the five year observation period. The birth year fixed effects in the main regression should take care of these level differences, but may be then entangled with an abnormally large or small cohort year. Column (2) therefore omits the birth cohorts born after 2004 and re-runs the main regression on the 1999-2004 cohorts. All of these six cohorts have a full five years of discharge data. The second trimester coefficient in Column (2) is actually larger than that of the main result in Column (1), strongly suggesting that the main result is robust to the inclusion of younger cohorts.

This Table shows that there is a large negative and highly statistically significant coefficient on the relative amount of sunlight during the second trimester, the exact same period we documented in Study I. Other trimesters show no effect. To interpret our second trimester coefficient, consider the following. Across this dataset, compared to the mean sunlight exposure within a specific county-month pair, the minimum and maximum levels are approximately -0.2 to 0.2 log points away. If everyone had the sunniest second trimester in the data set for their month and county of birth, this would thus yield on average approximately 20% more sunlight. That 20% increment would lower the rate of asthma emergency department discharges by 0.2*0.03 = 0.006, which at the mean rate of 0.10 observed in this dataset, is about a 6% reduction relative to the average. For a rough comparison, the 1989 introduction of two new surfactant drugs as a treatment for neonatal acute respiratory distress (considered a "home run" technology judged by cost effectiveness) reduced overall infant mortality by 5% (Chandra and Skinner, 2012). Thus, variation in sunlight even at the county-month level can exert a substantial effect.[24] These results pass robustness checks similar to those described for Study I (see the Appendix, Section 6, Table A.9).

Alternative Hypotheses: We provide evidence that within-location variation in sunlight across birth years affects the probability of asthma pathogenesis. For our results to provide evidence of a causal role for vitamin D, we must assume that our variation in

[24] If we repeat this analysis on a similar sample from Study I (babies born in December in New Jersey), we see a 4.8% decrease in asthma diagnoses at the mean. While these values are of similar magnitude, the two are in fact significantly different. This is not surprising though, as given the range of symptoms present in asthmatics, it seems likely that the effect of sunlight may differ for those who receive borderline diagnoses versus those with sufficiently extreme cases to entail emergency department visits.

sunlight correlates with maternal vitamin D levels and is uncorrelated with other factors that might actually determine asthma. As mentioned earlier, in Appendix 4 we provide suggestive evidence of the former. As an example of the latter, if sunnier weather causes pregnant women to smoke more, and smoking during the second trimester affects asthma incidence, we would see the same results in our data even though vitamin D would not be involved at all.

While we cannot rule out all such alternative stories, we wanted to address as many as we could given our data. First, at a broad level, we find no evidence in Study I (Table A.8) that the variation in sunlight we are using affects the probability of being low birth weight. If we interpret this outcome as a proxy for more general fetal health, then this provides evidence against any alternative channels that would affect overall fetal health as well (e.g., maternal exercise).

Second, from the fetal development literature, two specific stories that have been studied in particular involve smoking while pregnant and pollution.[25] Namely, there have been associational studies that document correlations between *in utero* exposure to each of maternal smoking and pollutants such as carbon monoxide and asthma later in life (e.g., Stick et al. 1996; Gilliland, Li, and Peters, 2001; Mortimer et al., 2008). We note that apart from only reporting correlations, the exact biological pathways *in utero* connecting these to asthma remain unclear. In comparison, the connection between vitamin D and sunlight is very well documented, and the pre-existing evidence on *in utero* vitamin D and asthma has a clear, well researched biological pathway. To consider these further though, we provide evidence in Appendix 5 that our variation in sunlight

[25] The negative effects of cigarette smoke on current asthmatics is very well documented; note here we are referring to the smaller literature on *in utero* exposure to smoke.

does not correlate with smoking (Table A.4), nor does adding controls for carbon monoxide levels *in utero* substantially diminish our results (Table A.5).

V. Conclusion

Our results from two completely independent studies provide evidence that greater sunlight during the second trimester of pregnancy reduces the probability the unborn child will later develop asthma. The presumed causal channel operates through the mother's conveyance of vitamin D to the fetus, the vast majority of which comes from her sunlight exposure. Similarly, we have provided evidence that sunlight exposure is correlated with the variation in weather we exploit. To judge the overall significance of our results, we note if the true effect were zero, the probability of finding effects at less than the 1% significance level in two independent studies is less than one in ten thousand.

Our results suggest low levels of maternal vitamin D during pregnancy may be more harmful than previously thought. As mentioned earlier, vitamin D deficiency and insufficiency are extremely widespread, with multiple studies of pregnant women documenting overall rates in excess of 80% (Bodnar et al., 2007; Holmes et al., 2009; Johnson et al., 2011). Hence, our results suggest that if we could raise vitamin D levels in these women, the savings in terms of both future quality of life and healthcare costs could be substantial.

The most practical methods for augmenting vitamin D levels would either be supplements or increased sunlight exposure. Both may be valid, but when regularly available, we prefer the latter for two reasons. First, many women already take supplements during pregnancy that include vitamin D, but some studies have found their insufficiency rates are similar to those who do not take supplements (e.g., Holmes et al.,

2009, or see Vieth et al., 2001). This may be due to inadequate dosages, absorption problems, or compliance issues, but it suggests supplementation is not completely straightforward. Second, sunlight exposure is a very efficient means of vitamin D production -- a rule-of-thumb in the medical literature is that most people may obtain a sufficient daily dose of vitamin D from 10 minutes of direct sunlight exposure (Holick, 2007).[26] Further, unlike supplements, which can only deliver vitamin D in fixed quantities, production of vitamin D from the skin automatically shuts off when levels are sufficiently high (Holick and Garabedian, 2006). The main cost of additional sunlight exposure would be a slightly higher risk of sunburn and melanoma, but given how little sun exposure most people need to obtain sufficient vitamin D compared to how much is necessary for a sunburn, this risk seems modest.

Additionally, while this study has focused on asthma due to data limitations, other diseases are linked to low levels vitamin D (e.g. multiple sclerosis, schizophrenia, and cardiovascular disease). Future work could consider whether increased sunlight reduces incidence of those diseases as well, perhaps using this paper's empirical framework.

Prior research has shown the mechanisms that underlie asthma are numerous and complex. We find evidence suggesting that sunlight during the second trimester (which we posit proxies for maternal vitamin D levels) is a significant factor. Given the extent of the asthma epidemic afflicting the world today, arriving at a better understanding of this complex disease and its numerous contributing factors is extremely important. This paper demonstrates that statistical analyses of existing data can provide an inexpensive complement to epidemiological and experimental studies in making such determinations.

[26] The specific amount required varies with ethnicity, altitude, latitude, etc. - there are several calculators available online, such as http://nadir.nilu.no/~olaeng/fastrt/VitD-ez_quartMEDandMED_v2.html.

Appendix

Table of contents:

1. Mean Comparison for Study I

As mentioned in the Data section, here we show that the means of our variables of interest are similar if we include our entire sample versus if we restrict to those for whom we have asthma data. This is simply to suggest that by ascertaining asthma status we were not biasing our sample in any way. We also note that the data we are using is used by the federal government for many of its asthma statistics, so others have had a vested interest in making sure these samples are similar.

Table A.1. Comparison of summary statistics across samples.

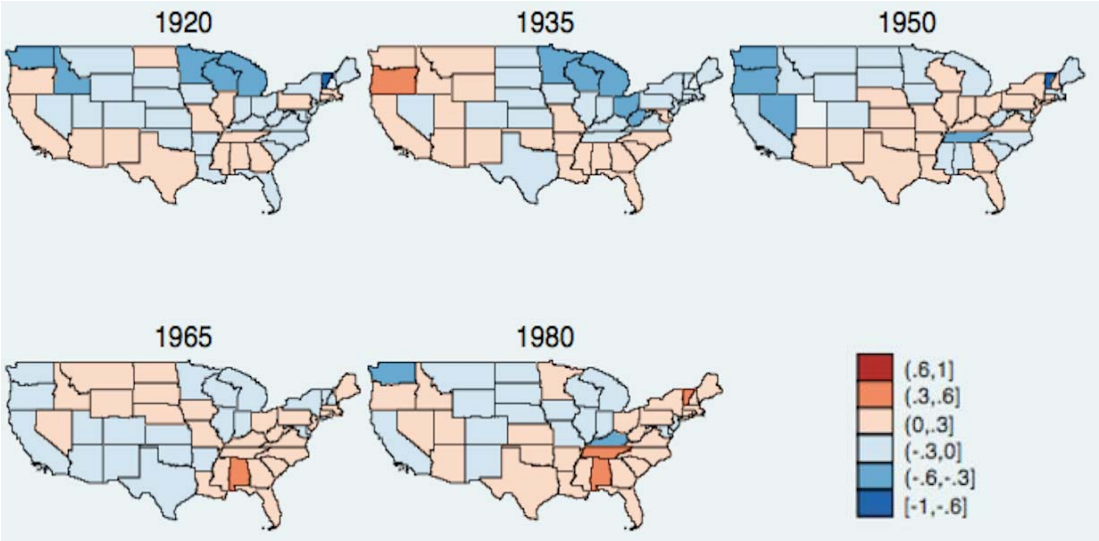
	Full Sample	Restricted Sample
N	595,093	264,701
Percent asthmatic:	11.49%	11.49%
Average year of birth:	1957.24	1956.06
Gender:		
Female	52.9%	56.0%
Major Ethnicities:		
White	81.6%	81.2%
Black/African American	15.0%	15.7%
Other	3.2%	2.8%
Hours of Sunlight:		
Mean ± Std. Dev.		
1st Trimester:	674.85±197.18	673.74±196.79
2nd Trimester:	677.52±197.33	676.58±197.28
3rd Trimester:	669.28±197.29	668.82±197.24

Notes: Restricted Sample refers to the full dataset restricted to those for whom we have asthma data on. (Note the 11.49% in both is mechanical due to the missing data.)

2. Sunlight Variation Figures

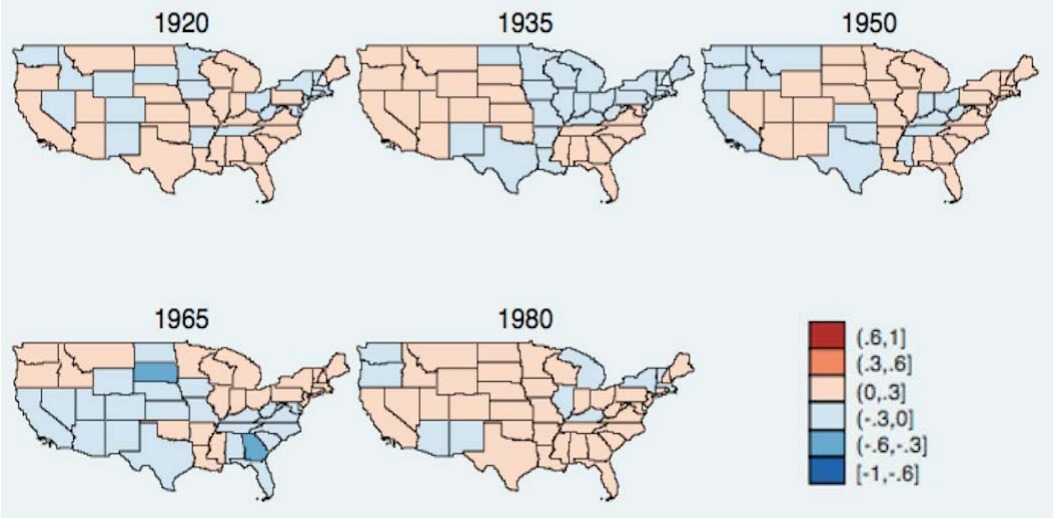
To better convey a sense of the variation in sunlight we are exploiting in each study, we here provide graphics that look at sunlight levels in different locations across years for two sample months (December and June).

F**igure A.1**. December sunlight variation, Study I



Notes: To illustrate the sunlight variation we exploit in Study I, this Figure demonstrates for select years the log deviations from the state-month mean level. In other words, we can see that December of 1980 was a particularly cold month for Washington state but particularly hot for Vermont, compared to their average Decembers (data from NOAA).

F**igure A.2**. June sunlight variation, Study I



Notes: Here we redo Figure A.1 but for the month of June (data from NOAA).

Figure A.3. Arizona sunlight variation, Study II

Figure A.3a. December weather

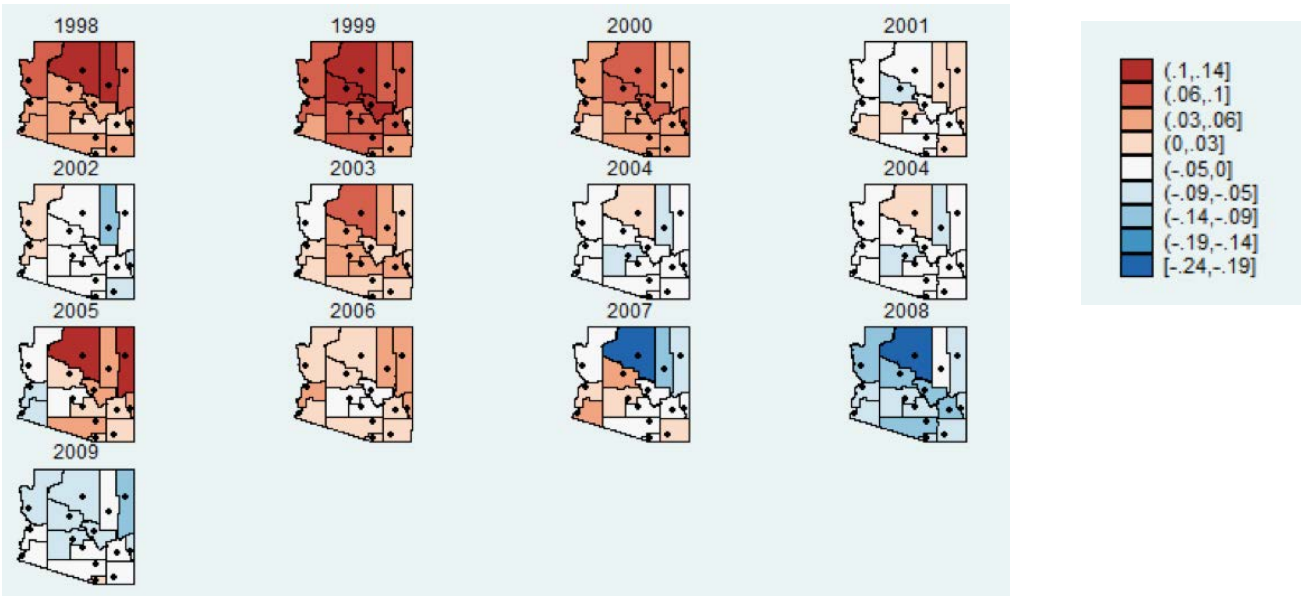
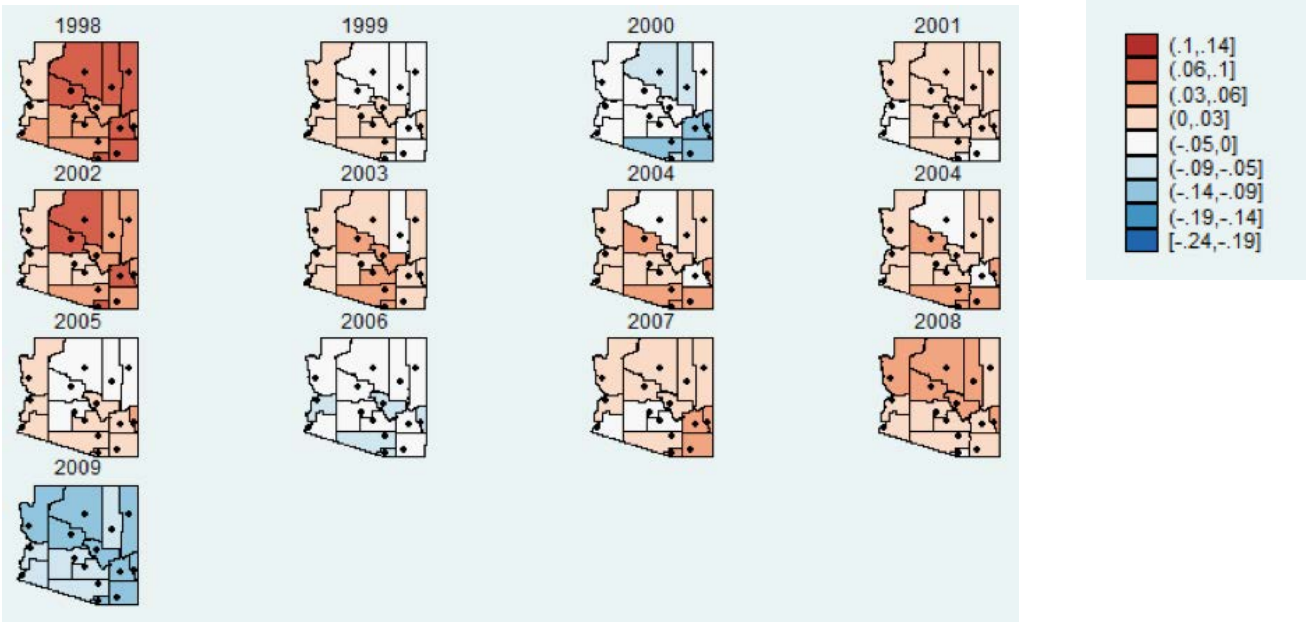


Figure A.3b. June weather



Notes: These two subfigures show the log deviations from the county-month mean for Arizona for December and June for the different years used in our analysis (data from NSRD). Scale ranges from -0.24 to 0.14 log points. The black dots indicate the population weighted county centroid.

Figure A.4. New Jersey sunlight variation, Study II

Figure A.4a. December weather

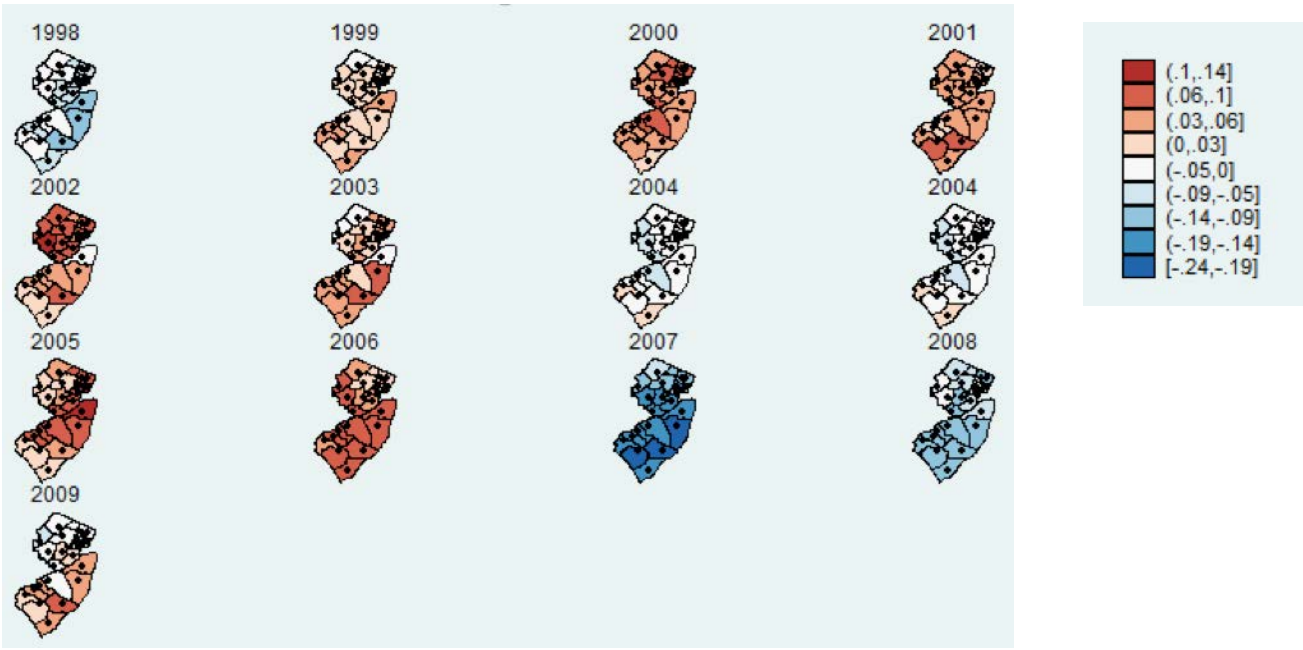
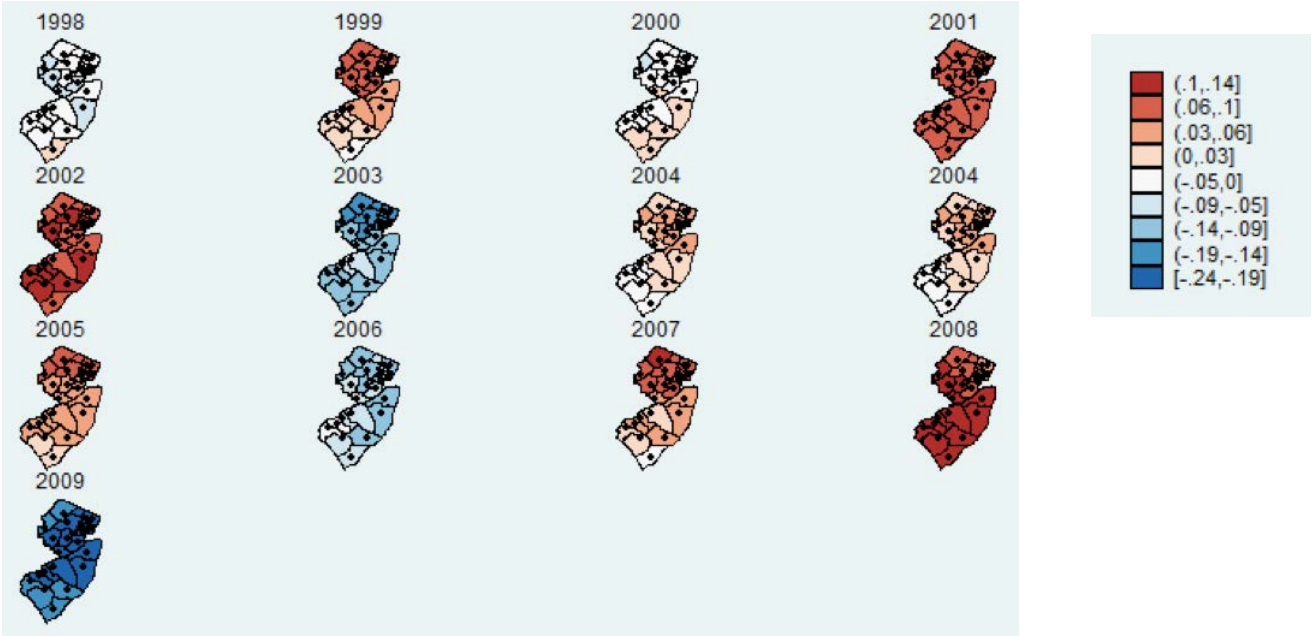


Figure A.4b. June weather



Notes: Here we repeat Figure A.3 with data from New Jersey (data from NSRD).

3. Month Level Estimation

Since data from both studies are available at the month level, we now rerun our main regressions at this level to see if we can provide evidence for when in the late pseudoglandular and canalicular stages vitamin D may be most important. Table A.2 contains the output from these regressions, where Column (1) is the month level version of Column (3) from Table III in the main text and Column (2) corresponds to Column (1) from Table IV.

These results show that no months outside our window of interest are significant in either dataset, but also that the main driver of the significance we see in the second trimester is from month five before the month of birth in both datasets. Thus, within the time frame we first posited, we find evidence that vitamin D may be particularly important during roughly weeks 10-18. As an example of how to interpret this magnitude, the Study I result predicts that doubling the total hours of sunlight during this window alone would reduce the mean incidence of asthma by about 4.8%.

Table A.2. Month level regressions.

Dependent variable:	(1) Asthma	(2) Asthma
Study number	I	II
Log Sunlight Month 1	-0.00260	0.01184
	(0.00530)	(0.01122)
Log Sunlight Month 2	0.00005	-0.00371
	(0.00430)	(0.01042)
Log Sunlight Month 3	0.00239	0.00910
	(0.00410)	(0.00768)
Log Sunlight Month 4	-0.00057	-0.00785
	(0.00314)	(0.00460)
Log Sunlight Month 5	-0.00794***	-0.01579***
	(0.00329)	(0.00557)
Log Sunlight Month 6	-0.00570	-0.00687
	(0.00269)	(0.00632)
Log Sunlight Month 7	-0.00100	-0.00507
	(0.00384)	(0.00662)
Log Sunlight Month 8	-0.00221	-0.01298
	(0.00404)	(0.01005)
Log Sunlight Month 9	0.00527	0.01785
	(0.00483)	(0.01207)
Location of Birth*Month of Birth Fixed Effects	x	x
Year of Birth Fixed Effects	x	x
Race Controls	x	
Gender	x	
N	264,524	3,036
R^2	0.0108	0.7629

Notes: Here we repeat our main analysis at the finer month level to see if we can narrow down when during pregnancy vitamin D may be most important. Coefficients are reported for each study's respective measurement of sunlight exposure, where, for example, "Log Sunlight Month 9" refers to the sunlight present nine months prior to the birth month. In both datasets, we see no months outside our window of interest are significant and specifically Month 5 is highly significant in both (p=0.0048, 0.0011, respectively). Column (2) is run on birth years 1999-2009.

4. Sunlight and Time Spent Outdoors

For our analysis to provide evidence between a possible connection between maternal vitamin D levels and asthma incidence, it is critical that the variation in sunlight we are considering correlates with actual vitamin D levels in pregnant women. In other words, fixing the location and time of year, if the weather one year is particularly sunny, do the people there tend to have higher vitamin D levels? Given that seasonal variation in vitamin D levels is well documented, the answer to this question may seem clear, but note that it is a slightly different variation in sunlight that we are exploiting.

Even if exposure to sunlight leads to vitamin D production, our hypothesis relating sunny weather *in utero* to asthma protection would fail if sunnier days led people, for example, to get less exposure to the sun, say by spending most of their time inside with air conditioning. To rule out this possibility, we investigate whether fixing the location and time of year, more sunlight leads individuals to spend more time inside using data from the American Time Use Survey (ATUS).[27]

The ATUS, which began in 2003, asks individuals for a specific day to diary all of the activities that they did that day, and identifies the state of residence. Since the satellite sunlight data used in Study II overlaps the time period of the ATUS (i.e. 2003-2009) and is granular to the day level, we can investigate whether individuals spend relatively more time on outdoor activities on relatively sunnier day.

[27] If individuals still spend the same amount of time outside, and there is more sunlight during the hours they are outside, they would still be getting more vitamin D. Hence, even obtaining no effects of increased sunlight on time outside may not discredit our first stage.

Table A.3 shows the results from the regression of the total number of minutes spend on outdoor activities[28] on the log of average sunlight intensity on the reference day.

Table A.3. Time spent outdoors.

Dependent variable:	(1) Time spent on outdoor activities	(2) Time spent on outdoor activities	(3) Time spent on outdoor activities	(4) Time spent on outdoor activities
Log Average Sunlight Intensity for Reference Day	4.3997*** (0.03582)	4.3670*** (0.03550)	2.0065*** (0.03874)	2.0513*** (0.03769)
Reference Day of the Week Fixed Effect		x	x	x
Year and State-Day-of-Year Fixed Effects			x	x
Demographic controls				x
N	98,558	98,558	98,558	98,558
R^2	0.0040	0.0063	0.2451	0.2537

Notes: Regression output of time spent on outdoor activities on average sunlight for the population-weighted state centroid on the reference day. The results indicate that greater sunlight results in more time spent outdoors ($p<10^{-6}$ for all). All regressions are done with robust standard errors clustered at the state level.

Column (2) adds fixed effects for the days of the week (e.g., Saturday), since individuals generally have more time for leisure on the weekends. Column (3), analogous to the main results above, adds day of year – state fixed effects (e.g. July 9 in Kansas) to control for time invariant local tastes and cultures. Finally, Column (4) adds demographic

[28] Activities that are coded as "outdoor" include playing baseball, biking, boating, climbing, spelunking, caving, participating in equestrian sports, fishing, playing football, golfing, hiking, playing hockey, hunting, playing racquet sports, participating in rodeo competitions, rollerblading, playing rugby, running, skiing, ice skating, snowboarding, playing soccer, softball, playing volleyball, walking, and participating in water sports.

controls. The coefficient in Column (4) can be interpreted as follows. A sunny day might have three times the average sunlight intensity from a cloudy day, or a one log-point increase. Across the population, the average effect is to spend 2 more minutes on outdoor activities. But given that 90% of the population spends 0 minutes on outdoor activities, the size of the effect on those who spend time outdoors could be scaled up to approximately 20 minutes, which is a more plausible intensive margin. This shows that individuals are reacting to increased sunlight by spending more time outdoors, suggesting that the first stage of our identification strategy is sound.[29]

[29] A Tobit regression, which accounts for the censored individuals (i.e. those who spend no time outside), confirms this back-of-the-envelope calculation and returns a coefficient (using outdoor time and sunlight intensity demeaned at the day-state level) of 22.6^{***} (3.05), which is approximately 10 times (i.e. 1/(1-90%)) the 2.05 coefficient in Column (4).

5. Sunlight and Smoking, Pollution

In this section, we explore whether our variation in sunlight might actually be picking up two factors other than vitamin D which may be affecting asthma rates.

Smoking: One could imagine various stories by which the variation in sunlight we are exploiting correlates with smoking rates. For example, since a sunny day can reduce stress and gives many individuals more options for leisure and amusement, one might expect it to reduce smoking rates. On the other hand, individuals now mostly smoke outdoors (as opposed to in their workplaces or homes), and so a less sunny day could result in reduced smoking. If we believe smoking could impact fetal lung development and therefore asthma rates, then our observed sunlight coefficient may just be picking up an effect from smoking behavior.

The ATUS, used above, also asked about the number of minutes per day spent smoking.[30] This variable can be used to test the relative impact of sunlight on smoking, using the same econometric framework as above. Table A.4 shows the results of this analysis.

[30] The ATUS has a single code for "Tobacco and drug use." We are assuming here that since tobacco use rates are much higher than drug use rates and that smoking is still the primary method of tobacco use this variable can safety used to represent time spent smoking.

Table A.4. Study II, Smoking results.

Dependent variable:	(1) Time spent on smoking	(2) Time spent on smoking	(3) Time spent on smoking	(4) Time spent on smoking
Log Average Sunlight Intensity for Reference Day	-0.04423 (0.04540)	-0.04390 (0.04567)	-0.05950 (0.08338)	-0.05889 (0.08253)
Reference Day of the Week Fixed Effect		x	x	x
Year and State-Day-of-Year Fixed Effects			x	x
Demographic controls				x
N	98,558	98,558	98,558	98,558
R^2	0.0000	0.0006	0.4061	0.4073

Notes: Regression output of time spent on smoking and drug use on average sunlight for the population-weighted state centroid on the reference day. The results indicate no relationship between greater sunlight and time spent on smoking / drug use (p>0.1 for all). All regressions are done with robust standard errors clustered at the state level.

The coefficients are extremely small in magnitude and are statistically insignificant. A one log-point increase would increase the mean amount of time spent smoking by only a few seconds, which is even less than one cigarette. With the standard errors in Column (4), we can rule out at the 5% level any effect bigger than 13 fewer seconds ((-0.05 -0.08*2)*60) or 6 more seconds spent smoking on an unusually sunny day.[31]

[31] One might be concerned that the linear model used here is mis-specified, since the vast majority of individuals in the sample (98.5%) spend 0 minutes on tobacco or drug use. Using a Tobit regression (as above) of using smoking time on sunlight intensity, both demeaned at the day-state level, gives a coefficient of -1.78 (1.67). While this is larger in magnitude, is it is still far from being statistically significant.

Pollution: One may be also be concerned about a potential correlation between sunlight and pollution levels. The study cited in the main text specifically refers to carbon monoxide (CO), which in a separate literature has consistently been shown to be a pollutant with adverse effects on child health (Neidell 2004; Currie, Neidell, and Schmieder, 2009; Schlenker, Walker 2011).

To show that our results are robust to this potential issue, we apply the regression framework of Study II to pollution data,[32] and also to a combined regression that incorporates both pollution and sunlight data. The analysis below uses measured CO level (in parts per million), averaged over each calendar month and year. As with the satellite sunlight data, each county is matched with the pollution monitor that is the closest to its population weighted centroid.

The log of a three month moving average level of CO is somewhat negatively correlated with sunlight (-0.2). This adds plausibility to the concern described above, since higher *in utero* sunlight levels, which should reduce asthma rates, would be coincident with lower *in utero* CO levels, which could also reduce asthma rates.

In Table A.5, Column (1) repeats that main results for Study II from above for comparison. Column (2) replaces the log of average sunlight in each trimester with the log of the average CO level in each trimester. Finally, Column (3) includes both sets of independent variables to see how much CO levels reduce the magnitude and significance of this paper's main finding.

[32] Daily pollution data from the EPA is publically available at http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html (see Currie, Neidell, and Schmieder, 2009 and Sanders, 2012)

Table A.5. Study II, Carbon Monoxide results.

Dependent variable:	(1) Asthma	(2) Asthma	(3) Asthma
Log of Average Sunlight intensity 3rd Trimester	0.0114		0.0118
	(0.0244)		(0.0241)
Log Average Sunlight intensity 2nd Trimester	-0.0319***		-0.0287***
	(0.0098)		(0.0092)
Log Average Sunlight intensity 1st Trimester	0.0205		0.0254
	(0.0197)		(0.0174)
Log Average CO level 3rd Trimester		0.0092	0.0094
		(0.0090)	(0.0088)
Log Average CO level 2nd Trimester		-0.0010	-0.0005
		(0.0072)	(0.0072)
Log Average CO level 1st Trimester		0.0056	0.0051
		(0.0067)	(0.0067)
County of Birth*Month of Birth Fixed Effects	x	x	x
Year of Birth Fixed Effects	x	x	x
Birth years	1999-2009	1999-2009	1999-2009
N	3,036	3,036	3,036
R^2	0.7624	0.7628	0.7633

Notes: Here we repeat the analysis done earlier for Study II. Column (1) contains the same main result as above. Column (2) repeats above analysis using instead the log of the average CO level. Column (3) runs a "horse race" including both sets of independent variables. Robust standard errors clustered at the county level. The *p*-value for the second trimester sunlight coefficient in Column (3) is 0.005, versus 0.004 in Column (1).

.

Column (3) shows the results of including both *in utero* CO level and sunlight levels. The second trimester sunlight result is still extremely statistically significant, but is now about 10% smaller in magnitude. While this is consistent with a hypothesis that *in utero* CO levels do have an effect on asthma, and that this effect is somewhat correlated

with the effect of sunlight, it does not diminish the robustness of the main result of this paper. The effect of sunlight is not simply proxying for effects of pollution, but rather represents its own substantial result.

6. Robustness checks

In this section, we detail several robustness checks mentioned in the main text. First, we re-run our analysis for both studies with a linear independent variable instead of a log transformed one. Second, for both studies, we first add an additional pregnancy's worth of trimesters prior to conception to our regression to confirm that sunlight in those periods does not relate to asthma incidence. Third, we then add another pregnancy's worth of trimesters post-birth to confirm that those as well are not related to our dependent variable of interest. The second trimester variable retains its significance in all of these regressions in both datasets.

Finally, for Study I we repeat our analysis using two dummy variables which we do not hypothesize are related to sunshine exposure (i.e., low birth weight, defined as weighing less than 5.5 pounds at birth, and whether or not someone's father graduated high school) to show that we find no evidence that sunlight has an effect on these outcomes. We note that the NHIS unfortunately had fewer data points on these last two measurements, but we chose them since other NHIS variables describe post-birth characteristics which may be endogenous. Data limitations prevent us from looking at these two variables for Study II, so instead we use the share of white and share of males in each birth month-year-county cohort as dependent variables. We find no evidence that sunlight has an effect on these outcomes.

Table A.6. Regression output from linear specification, Study I

Dependent variable:	(1) Asthma	(2) Asthma
Independent variable is:	Log	Linear
Log Hours of Sun 3rd Trimester	0.00140	0.0000088
	(0.00711)	(0.00001104)
Log Hours of Sun 2nd Trimester	-0.01653***	-0.0000238**
	(0.00596)	(0.00000927)
Log Hours of Sun 1st Trimester	0.00360	0.0000014
	(0.00717)	(0.00001131)
State of Birth*Month of Birth Fixed Effects	x	x
Year of Birth Fixed Effects	x	x
Race controls	x	x
Gender	x	x
N	264,533	264,533
R^2	0.01073	0.01073

Notes: Column (1) contains the main analysis from Study I, whereas Column (2) repeats the analysis with linear sunlight terms instead of log transformed ones (for the second trimester terms, p= 0.0079 and 0.0133, respectively). Robust standard errors clustered at the state level.

Table A.7. Regression output from linear specification, Study II

Dependent variable:	(1) Asthma	(2) Asthma
Independent variable is	Log	Linear
Average Sunlight intensity 3rd Trimester	0.0114 (0.0244)	0.0000765 (0.000116)
Average Sunlight intensity 2nd Trimester	-0.0319*** (0.0098)	-0.000180*** (0.0000549)
Average Sunlight intensity 1st Trimester	0.0205 (0.0197)	0.0000991 (0.0000962)
County of Birth*Month of Birth Fixed Effects	x	x
Year of Birth Fixed Effects	x	x
Birth years	1999-2009	1999-2009
N	3,036	3,036
R^2	0.7624	0.7624

Notes: Column (1) contains the main analysis from Study II, whereas Column (2) repeats the analysis with linear sunlight terms instead of log transformed ones. The second trimester variable is highly significant in both, with p=0.004 and 0.004, respectively. Robust standard errors clustered at the county level.

Table A.8. Regression output from robustness checks, Study I.

Dependent variable:	(1) Asthma	(2) Asthma	(3) Low Birth Weight	(4) Dad HS Grad
Log Hours of Sun 3rd Trimester	0.00140	0.00190	-0.00020	0.01302
	(0.00684)	(0.00684)	(0.03254)	(0.03217)
Log Hours of Sun 2nd Trimester	-0.01676***	-0.016434**	0.00618	0.03763
	(0.00612)	(0.00631)	(0.02509)	(0.03993)
Log Hours of Sun 1st Trimester	0.00336	0.00482	-0.01258	0.03136
	(0.00736)	(0.00737)	(0.02854)	(0.04427)
Log Hours of Sun 1st Trimester, Pre-conception	0.001289	0.00281		
	(0.00656)	(0.00638)		
Log Hours of Sun 2nd Trimester, Pre-conception	-0.000556	-0.00007		
	(0.00659)	(0.00656)		
Log Hours of Sun 3rd Trimester, Pre-conception	0.00147	0.00196		
	(0.00658)	(0.00666)		
Log Hours of Sun 1st Trimester, Post-birth		-0.00650		
		(0.00547)		
Log Hours of Sun 2nd Trimester, Post-birth		-0.00887		
		(0.00814)		
Log Hours of Sun 3rd Trimester, Post-birth		0.00160		
		(0.00716)		
State of Birth*Month of Birth Fixed Effects	x	x	x	x
Year of Birth Fixed Effects	x	x	x	x
Race Controls	x	x	x	x
Gender	x	x	x	x
N	264,394	264,252	18,787	19,646
R^2	0.0108	0.0108	0.0416	0.0512

Notes: Regression output from robustness checks. Here we provide evidence that the sunlight specifically during the late pseudoglandular and canalicular stages is important by showing that the sunlight levels both before and after the pregnancy are insignificant, while our main coefficient of interest retains its significance (p=0.0086, 0.0121). We also repeat our analysis on low birth weight and whether or not the individual's father was a high school graduate, variables which are not hypothesized to be related to vitamin D *in utero* to verify that we see no effect. Robust standard errors are clustered at the state level.

Table A.9. Regression output from robustness checks, Study II.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Asthma	Asthma	Asthma	Share of births male	Share of births male	Share of births white	Share of births white
Log Average Sunlight intensity 3rd Trimester	-0.00902	-0.00389	0.00258	-0.00125	-0.00203	-0.00837	-0.00459
	(0.02607)	(0.02645)	(0.02809)	(0.0119)	(0.0207)	(0.0.0136)	(0.0198)
Log Average Sunlight intensity 2nd Trimester	-0.04995***	-0.04830***	-0.04102***	-0.0264*	-0.0171	-0.0127	-0.0151
	(0.01249)	(0.01265)	(0.01451)	(0.0131)	(0.0149)	(0.00901)	(0.00147)
Log Average Sunlight intensity 1st Trimester	0.00332	-0.00452	0.00094	0.00338	0.00738	-0.0237	-0.0149
	(0.01942)	(0.02131)	(0.02115)	(0.0112)	(0.0146)	(0.0124)	(0.0175)
Log Average Sunlight intensity 1st Trimester, Pre-conception		0.01989	0.02319				
		(0.02259)	(0.02374)				
Log Average Sunlight intensity 2nd Trimester, Pre-conception		0.00137	0.01002				
		(0.02064)	(0.00252)				
Log Average Sunlight intensity 3rd Trimester, Pre-conception		-0.02983	-0.02067				
		(0.02349)	(0.02473)				
Log Average Sunlight intensity 1st Trimester, Post-birth			0.03150				
			(0.02285)				
Log Average Sunlight intensity 2nd Trimester, Post-birth			0.01637				
			(0.01518)				
Log Average Sunlight intensity 3rd Trimester, Post-birth			0.00974				
			(0.01721)				
Birth Years	2000-2008	2000-2008	2000-2008	1999-2009	1999-2004	1999-2009	1999-2004
County of Birth*Month of Birth Fixed Effects	x	x	x	x	x	x	x
Year of Birth Fixed Effects	x	x	x	x	x	x	x
N	2,484	2,484	2,484	3,036	1,656	3,036	1,656
R^2	0.7718	0.7723	0.7727	0.1061	0.1804	0.9702	0.9751

Notes: Here, we repeat the analysis done earlier for Study II. Whereas our results in the main text included individuals born in 1999-2009, in order to trace out additional sunlight data, we must restrict our sample to births in 2000-2008 here. Hence, the first column repeats the original analysis on this smaller sample. In all of the asthma regressions, the second trimester variable retains its significance, with p=0.001, 0.001, and 0.010, respectively. For placebo regressions from this dataset, we run on the share of male and white births in each birth month-county-year cohort (Columns 4-7). None of the coefficients are significant at the 5% level, with p=0.055, 0.264, 0.171, and 0.316 respectively. Robust standard errors clustered at the county level.

References:

Almond, Douglas, and Janet Currie. "Killing me softly: The fetal origins hypothesis." *The Journal of Economic Perspectives*, 25 (2011), 153-172.

Almond, Douglas, Lena Edlund, and Mårten Palme. "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden." *Quarterly Journal of Economics*, 124(4) (2009), 1729-1772.

Barnett, Sarah Beth L., and Tursynbek A. Nurmagambetov. "Costs of asthma in the United States: 2002-2007." *Journal of Allergy and Clinical Immunology,* 127(1) (2011), 145-152.

Belderbos, Mirjam E., et al. "Cord blood vitamin D deficiency is associated with respiratory syncytial virus bronchiolitis." *Pediatrics*, 127(6) (2011): e1513-e1520.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson. "Early life health interventions and academic achievement." *American Economic Review*, 103(5) (2013), 1862-1891.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics*, 122(1) (2007), 409-439.

Bleakley, Hoyt. "Disease and development: evidence from hookworm eradication in the American South." *Quarterly Journal of Economics*, 122(1) (2007), 73-117.

Bodnar, Lisa M., et al. "High prevalence of vitamin D insufficiency in black and white pregnant women residing in the northern United States and their neonates." *Journal of Nutrition*, 137(2) (2007), 447-452.

Bossé, Yohan, Karim Maghni, and Thomas J. Hudson. "1α, 25-Dihydroxy-vitamin D3 stimulation of bronchial smooth muscle cells induces autocrine, contractility, and remodeling processes." *Physiological Genomics*, 29(2) (2007), 161-168.

Bound, John, and David A. Jaeger. "Do compulsory school attendance laws alone explain the association between quarter of birth and earnings?" *Research in Labor Economics*, 19 (2001), 83-108.

Brun, P., et al. "Vitamin D-dependent calcium-binding proteins (CaBPs) in human fetuses: comparative distribution of 9K CaBP mRNA and 28K CaBP during development." *Pediatric Research*, 21(4) (1987), 362-367.

Buckles, Kasey S., and Daniel M. Hungerman. "Season of birth and later outcomes: Old questions, new answers." *Review of Economics and Statistics*, 95(3) (2013), 711-724.

Burri, Peter H. "Fetal and postnatal development of the lung." *Annual Review of Physiology*, 46(1) (1984), 617-628.

Camargo, Carlos A., et al. "Maternal intake of vitamin D during pregnancy and risk of recurrent wheeze in children at 3 y of age." *American Journal of Clinical Nutrition*, 85(3) (2007), 788-795.

Castro, Mario, et al. "Effect of Vitamin D3 on Asthma Treatment Failures in Adults With Symptomatic Asthma and Lower Vitamin D Levels The VIDA Randomized Clinical Trial." *Journal of the American Medical Association* 311(20) (2014), 2083-2091

Chandra, Amitabh and Jonathan Skinner. "Technology Growth and Expenditure Growth in Health Care." *Journal of Economic Literature*, 50(3) (2012), 645–680

Chen, Tai C. "Photobiology of vitamin D." In *Vitamin D* Ed. MF Holick. Humana press (1999) 17-37.

Cruz, Alvaro A., et al. *Global surveillance, prevention and control of chronic respiratory diseases: a comprehensive approach*. Eds. Jean Bousquet, and N. G. Khaltaev. (Geneva, World Health Organization, 2007).

Currie, Janet, and Matthew Neidell. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics*, 120(3) (2005), 1003-1030.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. "Air pollution and infant health: Lessons from New Jersey." *Journal of Health Economics* 28(3) (2009), 688-703.

Currie, Janet, and Hannes Schwandt. "Within-mother analysis of seasonal patterns in health at birth." *Proceedings of the National Academy of Sciences*, 110(30) (2013), 12265-12270.

Devereux, Graham, et al. "Maternal vitamin D intake during pregnancy and early childhood wheezing." *American Journal of Clinical Nutrition*, 85(3) (2007), 853-859.

Disanto, Giulio, et al. "Month of Birth and Thymic Output." *Journal of the American Medical Association, Neurology*, 10(4) (2013), 527-528.

Erkkola, M., et al. "Maternal vitamin D intake during pregnancy is inversely associated with asthma and allergic rhinitis in 5-year-old children." *Clinical & Experimental Allergy*, 39(6) (2009), 875-882.

Fletcher, Jason M., Jeremy C. Green, and Matthew J. Neidell. "Long term effects of childhood asthma on adult health." *Journal of Health Economics*, 29(3) (2010), 377-387.

Gale, Catharine R., et al. "Maternal vitamin D status during pregnancy and child outcomes." *European Journal of Clinical Nutrition*, 62(1) (2007), 68-77.

Gaultier, C., et al. "Lung mechanics in rachitic rats." *American Review of Respiratory Disease*, 130(6) (1984), 1108-1110.

Gilliland, Frank D., Yu-Fen Li, and John M. Peters. "Effects of maternal smoking during pregnancy and environmental tobacco smoke on asthma and wheezing in children." *American Journal of Respiratory and Critical Care Medicine*, 163(2) (2001), 429-436.

Ginde, Adit A., Mark C. Liu, and Carlos A. Camargo. "Demographic differences and trends of vitamin D insufficiency in the US population, 1988-2004." *Archives of Internal Medicine*, 169(6) (2009), 626-632.

Glasgow, J. F., and P. S. Thomas. "Rachitic respiratory distress in small preterm infants." *Archives of Disease in Childhood*, 52(4) (1977), 268-273.

Heaney, Robert P., et al. "Human serum 25-hydroxycholecalciferol response to extended oral dosing with cholecalciferol." *American Journal of Clinical Nutrition*, 77(1) (2003), 204-210.

Heckman, James J. "Skill formation and the economics of investing in disadvantaged children." *Science* 312,5782 (2006), 1900-1902.

------, "The economics, technology, and neuroscience of human capability formation." *Proceedings of the National Academy of Sciences*, 104(33) (2007): 13250-13255.

Holick, Michael F. "Sunlight and vitamin D for bone health and prevention of autoimmune diseases, cancers, and cardiovascular disease." *The American Journal of Clinical Nutrition*, 80(6) (2004), 1678S-1688S.

------, "Vitamin D deficiency." *New England Journal of Medicine*, 357(3) (2007), 266-281.

Holick, Michael F., and Michele Garabedian. "Vitamin D: photobiology, metabolism, mechanism of action, and clinical applications." *Primer on the metabolic bone diseases and disorders of mineral metabolism.* 6th ed. *Washington, DC: American Society for Bone and Mineral Research,* (2006), 106-14..

Holmes, Valerie A., et al. "Vitamin D deficiency and insufficiency in pregnant women: a longitudinal study." *British Journal of Nutrition*, 102(06) (2009), 876-881.

Jensen, Robert. "Agricultural volatility and investments in children." *American Economic Review, Papers and Proceedings*, (2000), 399-404.

Johnson, Donna D., et al. "Vitamin D deficiency and insufficiency is common during pregnancy." *American Journal of Perinatology*, 28(01) (2011), 007-012.

Kassowitz, M."Tetanie and autointoxication in kindersalter." *Wien Med Presse*, 38(97) (1897), 139-141.

Kho, Alvin T., et al. "Vitamin D related genes in lung development and asthma pathogenesis." *BMC Medical Genomics*, 6(1) (2013), 47.

Lunghi, Barbara, et al. "1, 25-Dihydroxyvitamin D_3 inhibits proliferation of IMR-90 human fibroblasts and stimulates pyruvate kinase activity in confluent-phase cells." *Molecular and Cellular Endocrinology*, 115(2) (1995), 141-148.

Maccini, Sharon, and Dean Yang. "Under the weather: Health, schooling, and economic consequences of early-life rainfall." *American Economic Review*, (2009), 1006-1026.

Milet, M., et al. "The Burden of Asthma in California: A Surveillance Report." *Richmond, CA: California Department of Health Services, Environmental Health Investigations Branch* (2007).

Miyake, Y., et al. "Dairy food, calcium and vitamin D intake in pregnancy, and wheeze and eczema in infants." *European Respiratory Journal*, 35(6) (2010), 1228-1234.

Mortimer, Kathleen, et al. "Air pollution and pulmonary function in asthmatic children: effects of prenatal and lifetime exposures." *Epidemiology*, 19(4) (2008), 550-557.

National Center for Environmental Health, www.cdc.gov/Asthma/nhis/default.htm; last updated 6 March, 2014; accessed 10 August, 2014.

Neidell, Matthew J. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of Health Economics*, 23(6) (2004), 1209-1236.

Oster, Emily. "Witchcraft, weather and economic growth in renaissance Europe." *Journal of Economic Perspectives*, 18(1) (2004), 215-228.

Perez, R.; Ineichen, P.; Moore, K.; Kmiecikm, M.; Chain, C.; George, R.; and Vignola, F. "A New Operational Satellite-to-Irradiance Model." Solar Energy, 73(5) (2002):307–317.

Phokela, Sarabjit S., et al. "Regulation of human pulmonary surfactant protein gene expression by 1α, 25-dihydroxyvitamin D3." *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 289(4) (2005), L617-L626.

Sanders, Nicholas J. "What doesn't kill you makes you weaker prenatal pollution exposure and educational outcomes." *Journal of Human Resources*, 47(3) (2012): 826-850.

Schlenker, Wolfram, and W. Reed Walker. *Airports, air pollution, and contemporaneous health*. NBER Working Paper No. w17684, 2011.

Stick, Stephen M., et al. "Effects of maternal smoking during pregnancy and a family history of asthma on respiratory function in newborn infants." *The Lancet*, 348,9034 (1996), 1060-1064.

Stio, M., et al. "Vitamin D receptor in IMR-90 human fibroblasts and antiproliferative effect of 1, 25-dihydroxyvitamin D3." *IUBMB Life*, 43(6) (1997), 1173-1181.

Vieth, R., et al. “Wintertime vitamin D insufficiency is common in young Canadian women, and their vitamin D intake does not prevent it.” *European Journal of Clinical Nutrition*. 55: 1091-1097.

Weiss, Scott T., and Augusto A. Litonjua. "The *in utero* effects of maternal vitamin D deficiency: how it results in asthma and other chronic diseases." *American Journal of Respiratory and Critical Care Medicine*, 183(10) (2011), 1286-1287.