The Cicada Song: the Impact of Insecticides on Infant Health and Long-term Outcomes

Charles A. Taylor*

September 2019

Abstract

This paper utilizes a peculiar ecological phenomenon, the mass emergence of cicadas in 13 and 17-year cycles, to identify the impact of insecticides on human health and development. I rely on the fact that cicadas only damage woody plants (e.g., apple trees), through egg laying in branches and subsequent nymph-feeding on roots, and not agricultural crops such as corn or soy. Using the natural temporal and geographic variation of cicada emergence, I show that a sharp increase in insecticide use coincides with cicada emergence in counties with high levels of tree crop production. This is followed by a jump in next-year infant mortality in these counties as well as other negative infant health impacts. Looking at long-term impacts, I find evidence of lower elementary school test scores and then higher dropout rates nearly two decades later among exposed cohorts. Finally, I exploit variation in groundwater well usage to find evidence that insecticide exposure occurs mainly through the contamination of water supplies. *JEL Codes: I10, Q10, Q53, Q57.*

1 Introduction

Farmers in the US spend $7.9 billion annually on pesticides (US EPA 2017). Modern pesticides, along with other technological advances in agriculture, have brought about significant increases in productivity. But concerns have long been raised about the potential negative environmental and health impacts of pesticides given their toxicity by design.

*School of International and Public Affairs, Columbia University

Since the high-profile federal ban of DDT in 1972, dozens of pesticides have been banned by the EPA on account of their potential risk to humans and the environment (Buffington and Mcdonald 2006).

This paper focuses on insecticides, the second-most used type of pesticide after herbicides. I utilize an ecological phenomenon, the emergence of periodical cicadas (*Magicicada septendecula*), as a source of quasi-exogenous temporal and spatial variation in the application of insecticides to identify a potential causal channel for the impact of insecticides on health. My identification strategy hinges on the unique fact that cicadas emerge as mass broods in the same locations every 13 or 17 years. For example, Thomas Jefferson described the 'great locust years' of Brood II cicadas that arrived every 17 years at his home in Monticello, Virginia (Jefferson 1944). This same brood still emerges on schedule at Monticello 250 years later, most recently during the summer of 2013.

I find a significant increase in insecticide use in years and in counties experiencing a cicada emergence. This impact, however, is limited to counties with a large proportion of woody crops like fruit trees, and not herbaceous row crops like corn and soy. This is because cicadas only damage woody plants. Nymphs feed on tree roots and adult cicadas lay their eggs in small branches.

Using apple trees as a proxy for woody crop intensity, I exploit the variation and compare treated counties (i.e., counties with high apple production in years of a cicada emergence) to untreated counties. In the treated counties, I find a corresponding increase in county-wide insecticide use and subsequent increase in next-year infant mortality of 0.3 deaths per thousand births (the current mean in the US is six deaths) following a cicada emergence. The birth impact extends into the second year as farmers continue applying higher levels of insecticide to control cicada nymph establishment. An investigation of the quarterly impacts aligns with the timing and patterns of insecticide usage by farmers. Treated counties also see lower Apgar scores, a measure of newborn health, and an increased probability of premature birth. There is also evidence of long-term negative impacts in the form of lower elementary school test scores and higher high school dropout rates among exposed cohorts.

The findings are robust across model specifications, timing, and measures of woody crop intensity other than apple trees, as well as several falsification and robustness tests. The results are in line with the more general correlation found between insecticide use and infant mortality at a national level—notwithstanding the influence of cicadas and land use.

Furthermore, I find some evidence that the effect of insecticides on health manifests through a water channel. Counties with higher reliance on well water, whose groundwater source can be easily contaminated by agricultural runoff, experience greater negative impacts.

The findings are likely generalizable outside of just agriculturally-intensive regions. Tree crops cover a relatively small portion of US counties (always less than 5% of county land area, generally far less than 1%), especially compared to row crops like soy and corn which can account for a majority of acreage in many counties. Baseline pesticide use is relatively low in most tree-intensive counties. These facts support the conclusion that moderate levels of pesticides, not just extreme exposures, can have impacts on human health and development. And since this analysis looks only at average effects at the county level, it likely understates the magnitude of health impacts among those living in close proximity to insecticide application.

Overall this paper contributes to the environmental and health economics literature on the health impacts of agricultural inputs. While acknowledging the large benefits of pesticides to agricultural productivity, the findings warrant caution in the over-application of insecticides. This paper also provides an example of how ecological phenomena like cicadas may be used to generate quasi-random variation that can be employed to answer important economic and public health questions.

2 Background

2.1 Pesticides and health

Pesticides, and insecticides in particular, are toxic by design. Many were initially developed for warfare purposes. One prominent insecticide type, organochlorides (e.g., DDT), opens sodium channels in the nerve cells; another, organophosphates, targets the nervous system like the nerve agents in chemical weapons.

While laboratory and controlled studies have documented the negative impacts of pesticides on organisms and ecosystem services such as water quality, few have demonstrated a direct causal link between pesticides and human health. Almond and Currie 2011 show that fetal shocks, particularly ones occurring early in a pregnancy, can have long-lasting impacts. Environmental shocks including heavy metal exposure, high temperatures, and air pollution have been causally linked to adverse birth outcomes (Chay and Greenstone 2003, Zheng et al. 2016).

But there is little evidence causally linking pesticides to health outcomes like infant mortality, low birth weight, premature birth, and birth abnormalities. And no study, to my knowledge, has directly linked pesticide exposure to long-term outcomes like educational achievement and attainment.

Most estimates of the health impacts of pesticides come from non-randomized studies with small sample sizes (Jurewicz et al. 2006, Andersson et al. 2014). Many focus on occupationally exposed groups like pesticide applicators who are unlikely to be representative of the broader population. Regidor et al. 2004 find higher levels of still births and infant deaths within 24 hours of birth, while Garry et al. 2002 document an increase in birth defects among farm families, especially for conceptions occurring during the spring pesticide application season. Bell et al. 2001 similarly highlights the impact of pesticide exposure during the first trimester. Winchester et al. 2009 find elevated levels of agricultural chemicals in water to be correlated with birth defects. Schreinemachers 2003 finds that birth defects increase with a county's wheat acreage, which they use as a proxy for herbicide exposure. Larsen et al. 2017 use detailed spatial and micro-level panel data in California to show that pesticide exposure increases adverse birth outcomes among populations exposed to high quantities of pesticides (i.e., 95$^{\text{th}}$ percentile exposure). Brainerd and Menon 2014 exploit variation in planting times to link agrichemical exposure to adverse birth outcomes in India. Rauh et al. 2012 find evidence of long-term impacts in the form of lower IQ scores among a small sample of children exposed to insecticides in utero.

This paper builds on the approach of Frank 2018, who exploits White Nose Syndrome in bat populations as a source of variation in insecticide application. Spread by an invasive fungus, White Nose Syndrome increases bat mortality in exposed geographies. And since bats act as a sort of natural insecticide as the eaters of insects, White Nose Syndrome results in the increased use of insecticides by farmers. This results in an increase in infant mortality that primarily affects female infants.

2.2 Cicadas and Insecticides

Periodical cicadas (*Magicicada septendecula*) occur throughout the eastern half of the US.[1] There are fifteen extant broods, three of which are on 13-year cycles and twelve of which are on 17-year cycles. Some counties receive two or more broods. Figure 1 is a map showing each brood's range, cycle, and next year of emergence.

There is ample agronomic and ecological research on cicadas and tree health, with a considerable focus on fruit trees in particular. Cicadas spend most of their lives underground feeding on the xylem fluids from the roots of deciduous trees before synchronously emerg-

[1] There are several species of annual (i.e., non-periodical) cicadas that exist globally, including in ranges that overlap with periodical cicadas in the US. But the populations of such species do not tend to vary greatly year to year.

ing in the spring at any given location. Emergence densities of 1.5 million cicadas per acre have been reported (Dybas and Davis 1962), representing some of the highest biomass values of any naturally occurring terrestrial creature. Cicadas remain active for four to six weeks to mate and lay their eggs in small tree branches (i.e., oviposition), causing harm especially to young trees. When the eggs hatch, the nymphs fall to the ground to begin their development. Tree growth is further damaged by cicada nymphs feeding on tree roots, which can reduce growth by up to 30% (Karban 1980).

Both processes have a negative impact on orchard trees. In an early study, Hamilton 1961 reported a complete loss among unprotected young apple and pear trees in the Hudson Valley following a cicada event in 1945. Karban 1982 conducted an experiment on apple trees and found that removing cicada nymphs significantly increased wood accumulation compared to trees with nymphs present.

Most commercial tree growers and serious gardeners are well aware of the damage that cicadas can cause, especially to young trees. Using insecticides to mitigate cicada damage is well documented. Hamilton 1961 describes the process and efficacy of spraying trees with insecticides to kill adult cicadas as well as soaking the soil with insecticides to control nymphs. Lloyd and White 1987 recommended killing off understory grasses to starve young nymphs. Penn State provides publicly-available online recommendations to tree fruit growers on cicada management, including the types of pesticides to use and application methods (Krawczyk 2017).

3 Data

3.1 Cicada data

The US Forest Service provides shapefiles with county-level presence-absence data on periodical cicadas by brood with emergences projected through 2031. The dataset was originally compiled from Koenig et al. 2011. Given the temporal and spatial consistency of cicada emergence, I extend the time series further into the past using each brood's 13 or 17-year cycle assuming that cicada emergence occurred in the same counties. While there are some examples of accelerations in cycles and changes in the range of broods (Williams and Simon 1995), cicada behavior has been remarkably consistent for the most part.

3.2 Agricultural data

The land use dataset was compiled from the USDA's National Agricultural Statistics Service (NASS) online tool and from the historical U.S. Census of Agriculture, available online through the Inter-university Consortium for Political and Social Research (Haines et al. 2014). I collected various measures of apple and woody crop intensity at the county-year level (i.e., number of trees, acres, production in bushels)[2]. I choose apples as my preferred explanatory variable because apples are the historically dominant tree crop in the US. There is also ample agronomic and ecological literature on the effect of cicadas on apple trees, as described earlier. Apple production is well-distributed geographically among US states, with top producers in the Northeast (NY, MA, CT), Central-Midwest (PA, MI, OH), and the South (VA, NC).

Unfortunately, an annual time series cannot be constructed for tree crop variables for sev-

[2] County-year data values of '(D)', which NASS uses to denote confidentiality, were coded as not available, and values of '(Z)', which denotes being too small to estimate, were coded as zero. Given that only positive values are included in NASS output, excluded county-years are assumed to have a value of zero

eral reasons: the agricultural census takes place every five years, variables were not measured consistently over time, and surveys in the 1970s and 1980s only included 50% of counties. Therefore I used a time invariant measure of county-level tree crop intensity, varying the base year as needed for analyses and robustness checks. All measures of agricultural intensity are standardized by county area.

3.3 Pesticide data

The US Geological Survey's National Water-Quality Assessment Project provides county-level pesticide use data from 1992 to 2016 (USGS 2019). Information is available by chemical constituent. My preferred measure is the sum of all insecticide-categorized constituents using the EPest-high measure in kilograms per county.[3] Insecticide intensity is also standardized by county area.

3.4 Infant health data

Infant mortality and birth outcome data come from the National Center for Health Statistics (NCHS 2019). NCHS Natality Data Files contain full records for data publicly available from 1968 to 1988, while records from 1989 to 2016 were obtained under confidentiality agreement. NCHS Linked Birth-Infant Death Data Files contain confidential microdata from 1995 to 2016. For longer-term analysis of infant mortality, I use ICPSR's County-Level Natality and Mortality Data, 1915-2007 (Bailey et al. 2016). The ICPSR data are averaged annually and do not allow for temporal or demographic disaggregation aside from race. I use ICPSR's preferred 'fixed' variables wherever available.

ICPSR's resident infant death data become available starting in 1941 and are based on the

[3] The categorized USGS data was kindly provided to me by Eyal Frank, who classified each of the tracked pesticide constituents by their reported function (i.e., insecticide, herbicide, fungicide), finding that 160 of the constituents had insecticidal properties.

residence county of the mother (rather than the county of birth occurrence). After 1988, ICPSR masks counties with populations less than 100,000, which presents challenges given that many of the counties of interest are agricultural and have populations lower than 100,000. Since the NCHS Linked Birth-Infant Death data begin in 1995, there is a data gap from 1989 to 1994 for low population counties. Starting in 1995 I use infant mortality rates derived from these linked files. I address concerns about sample composition by also running the analysis on subset of observations that ends in 1988.

I use the NCHS Linked Birth-Infant Death data from 1995 to 2016 to compute infant mortality rates at the sub-year level (i.e., quarter averages that can be linked to timing of insecticide application). I use NCHS Natality data from 1968 to 2016 to construct detailed birth outcome measures like Apgar scores, gestation time, and birth weight, as well as for constructing controls for maternal characteristics.

3.5 Education data

For educational achievement, I use standardized annual county-level test scores from the Stanford Education Data Archive (Reardon et al. 2018), available for the seven years from 2009 to 2015. I average across third, fourth, and fifth grade scores to produce an elementary school average score for each cicada exposure cohort (e.g., 3rd graders nine years after a cicada event, fourth graders ten years afterwards, and fifth graders eleven years afterwards).

For a measure of attainment, I construct a dataset on high school dropout rates using the National Center for Education Statistics Local Education Agency (School District) Universe Survey Dropout and Completion Data. I average across school districts to get county level values from 1991 to 2008. My preferred measure is twelfth grade dropout rate, which is the total number of twelfth graders dropping out of high school in a given year divided by the total number enrolled.

3.6 Water use data

County-level data on private well and groundwater use come from USGS's Estimated Use of Water in the United States. Estimates are available every five years from 1985 to 2015. I construct an average intensity measure over the years by dividing the number of people in a county that are reliant on private wells and/or domestic groundwater by the total number in the county.

4 Empirical Strategy

Cicada emergence is anticipated by both tree growers and, to a certain extent, the general population considering the ample news coverage leading up to 'cicada mania.' Figure 2 is a Google Trends chart of average monthly search volume for the word 'cicada' in metropolitan regions of Virginia, including Charlottesville, the area where Thomas Jefferson noted the creatures in his writings over two centuries ago. This event study demonstrates the distinct temporal pattern of periodical cicadas. The two distinct spikes in 2004 and 2013 coincide with the emergence years of the two endemic broods to the region.

Despite the public awareness, I argue that cicada emergence is effectively exogenous in relation to anything that could affect public health outcomes at a county level. Note that the Charlottesville region accounts for much of Virginia's fruit production, whereas Richmond and DC have few orchards. Yet in Figure 2 public interest in cicadas appears to follow similar patterns across regions. Cicada emergence therefore would act as a quasi-experiment where tree-intensive counties receive more insecticides during emergence years relative to the same counties during non-emergence years, and tree-intensive counties receive more insecticides relative to non-tree intensive counties in emergence years. I provide several robustness checks and alternative specifications to ensure that the exclusion restriction holds.

My empirical approach consists of two main steps: I use a triple difference estimator to first test whether there is an increase in insecticide use in treated counties in cicada emergence years, and second, whether there is a follow-on impact on infant health and longer-term outcomes. I restrict the sample to all the counties in the 32 states in the eastern half of the US that span the range of periodical cicadas.[4]

For the first step, I specify a model with insecticide use intensity, $insecticide_{it}$, as the dependent variable, measured in kilograms of insecticide per km^2 land area in county i in year t. The independent variable is a cicada presence-absence dummy, $cicada_{it}$, taking the value of 1 if there is a cicada emergence in county i in year t, and 0 otherwise. I interact this cicada dummy with a measure of tree crop intensity (e.g., apple production), $apple_i$, in county i which is unvarying over time:

$$insecticide_{it} = \beta_1 cicada_{it} + \beta_2 cicada_{it} * apple_i + \alpha_i + \gamma_{st} + \epsilon_{it} \tag{1}$$

where α_i includes county fixed effects and γ_{st} include state-year fixed effects. The former accounts for any time-invariant properties of the county that could affect outcomes and the latter flexibly controls for time-trends at the state level. The main coefficient of interest, therefore, is β_2, which estimates the change in insecticide use in tree crop-intensive counties driven by cicada emergence.

For health outcomes, I specify a similar model to Equation 1 but replace insecticide use intensity with infant mortality rate (infant deaths per thousand live births), $imr_{i,t+1}$, in county i in the following year, $t+1$:

$$imr_{i,t+1} = \beta_1 cicada_{it} + \beta_2 cicada_{it} * apple_i + \alpha_i + \gamma_{st} + \epsilon_{it} \tag{2}$$

The coefficient of interest is again β_2, which estimates the change in infant mortality rate

[4] Note that there are some counties in some states in which cicadas never emerge.

stemming from a cicada emergence in tree crop-intensive counties.

5 Results

5.1 Insecticides and Cicadas

The first analysis examines the relationship between insecticides and cicadas using the model specified in Equation 1. The sample is limited to the 26 years from 1992 to 2016 in which county-level USGS pesticide data exist. Table 1 regresses insecticide use on a cicada dummy and the cicada dummy interacted with fixed top-decile indicators (top 10^{th} percentile) for tree crop intensity. This subset of counties is well distributed across the country. For example, among the 242 counties in the top decile of apple producers in the eastern half of the US, 25 states have at least one county in this group. Model (1) shows the effect of cicada emergence on insecticide use alone. Model (2) interacts cicada emergence with the top decile of fruit acres, a broad category including non-citrus (apples, peaches, tree nuts, etc), citrus, and berry acres. Models (3)-(4) use the top decile of apple tree acres and apple bushels, respectively. Model (5) shows the results from Model (4) broken down by quantile.

Cicada emergence, in itself, is not associated with increased insecticide usage except in tree crop-intensive counties which see a strong increase in pesticide use in the range of 5-10 kg/km^2. This is a moderately large effect given that mean county pesticide use is 10 kg/km^2. Results are robust to varying the fixed effects (FE) included in the model.

Apple production is my preferred tree crop intensity measure going forward. Figure 3 shows the results of this regression after including various leads and lags of cicada emergence.[5] I see an increase in insecticide use only in the year of cicada emergence and the

[5] I limit the leads and lags to four years to limit the distortion of the event study from the fact that many counties receive more than one cicada brood, as seen in the national distribution map in Figure 1 and in

year following. This outcome aligns with the fact described earlier that farmers use insecticide to control the adult egg-laying population in the year of emergence and the cicada nymphs in the year that follows. There appears to be a moderate decline in insecticide use in later years, which can be explained by either reduced pest pressure following cicada emergence and chemical treatment, or a buildup of insecticide inventory by farmers.

As falsification tests, Table 2 shows that *only* insecticide use increases during cicada emergence, while herbicide and fungicide use do not change. This provides confidence that any resulting health impacts are attributable to insecticides and not a more general change in agricultural practices. Table 3 shows that cicada emergence is *not* associated with increased insecticide use in agriculturally-intensive counties containing a high proportion of row crops like corn and soy. This finding aligns with the fact that cicadas harm woody plants and not herbaceous annual crops.

5.2 Insecticide and Infant Mortality Correlation

After establishing the cicada-insecticide relationship, I examine the infant mortality impact. Infant mortality rate is measured as the number of infant deaths during the first year of life per thousand live births. Table 4 simply regresses infant mortality on insecticide use, ignoring any influence of cicadas or land use. The sample is restricted to the years 1992 to 2016 when both pesticide data and linked infant mortality data are available for all counties. Model (2) further restricts the sample to county-years with five or more births, and Model (3) includes only county-years that have at least one infant death. There is a small but imprecise positive relationship between pesticide use and next-year infant mortality.

Figure 4 shows the correlation results of Table 4 after adding annual leads and lags of in-

Virginia specifically in Figure 2.

secticide use to the regression.[6] I see the birth impact most clearly in the year following insecticide application, and there appears to be an uptick in deaths leading up to the current year's application. These results should be interpreted with caution given the possible omitted variables, autocorrelation in insecticide use, and the shortened time series.

Figure 5 shows the results by quarter. This time series is further limited to 1995-2016 when sub-annual Infant Birth/Death data is available. Note that 'q4' is the last quarter of the insecticide application year (i.e., October to December), and 'plus2_q4' is the last quarter of the second year following insecticide application. There is a peak in the positive association between insecticide use and infant mortality in the second quarter of the following year. This makes sense given that insecticide use tends to occur in late spring and summer for row crops and the fact that fetuses are highly susceptible to environmental shocks early in a pregnancy (Almond and Currie 2011).

Again, these correlations should be interpreted with caution given that the results are only significant after dropping county-year observations with zero infant deaths. All US counties and crop types are included in this analysis, and each can have different pesticide treatment practices, as well as potential pesticide exposure pathways. Further, these estimates could be biased given potential omitted variables and endogeneity concerns, as well as the limited time series beginning only in the 1990s.

5.3 Cicadas and Infant Mortality

To better address causality, I run a model specification based on Equation 2 in the Empirical Strategy section. Given the link established between insecticides and cicada emergence, as well as the correlation between insecticides and infant mortality, one would expect a relationship between cicada emergence and infant mortality if insecticides indeed have an

[6] Given the limited time series of insecticide data and since leads and lags reduce the number of observations, I limit this analysis to two leads and two lags. Results are noisier and less precise with the inclusion of additional years.

impact on health. In contrast to Table 4, this analysis allows for the use of a much longer time series. ICPSR starts tracking resident infant mortality at the county level in 1941, while USGS pesticide data is only available from 1992 to 2016. I restrict the sample to after 1950, which encompasses the post-WWII era when farmers started using synthetic pesticides at a large scale.[7]

Table 5 regresses next-year infant mortality on cicada emergence. Model (1) shows no significant impact of cicada emergence, in itself, on birth outcomes. Model (2) interacts cicada emergence with a dummy for high apple production (i.e., top decile counties). These counties experience an increase in infant mortality of 0.3 deaths per thousand. Models (3) and (4) use actual county-level apple production in bushels in 1964 and 1997, respectively. The 1964 measure of apple intensity helps address endogeneity concerns related to apple production changing over time.

A one standard deviation in apple production is equal to 170 bushels/km^2 in 1964 and 230 bushels/km^2 in 1997 on a cross-county basis. Therefore a one standard deviation increase in county apple production, when accompanied by cicada emergence, is associated with an increase in infant mortality of 0.1 deaths per thousand. General results hold when using log-values for infant mortality. Table 13 in the Appendix replicates this analysis by weighting the regression by the number of county-level births and keeping observations with less than five births. General results hold, though they are less precisely estimated, reflecting what may be a more rural phenomenon. Table 14 in the Appendix uses only observations from 1950 to 1988, which allows for a balanced panel.[8] The results hold with

[7] When analyzing birth outcomes, I drop counties with less than five births in a given year to minimize the inclusion of unreasonably high infant mortality rates due to small sample size (i.e., if there are two births in a county, and one death, IMR is 500 compared to the current US average of six). Results are robust to varying the birth cutoff threshold number. All health outcome standard errors are clustered at the state-level, which is the administrative level at which birth records are collected and aggregated. General results hold if standard errors are clustered at the county level.

[8] As discussed in the Data section, the ICPSR infant mortality dataset is limited after 1988 to counties with populations over 100,000, while the infant mortality rates derived from confidential NCHS Infant Linked Birth/Death files are not available until 1995. In general, the results hold whether I use the infant mortality dataset that I constructed that combines Linked Infant Birth/Death Files with historical ICPSR calculations, or if I use the ICPSR dataset which underwent some additional data cleaning as

slightly larger coefficients.

Figure 6 plots the cicada-apple interaction coefficients from Model (4) of Table 5 with the additional inclusion of cicada emergence leads and lags.[9] Infant mortality increases only in the year following cicada emergence, and there is a noisier indication that this impact may last through the second year. This matches well with Figure 3, which shows a two-year increase in pesticide use that can be explained by the fact that farmers control for both the cicada adult population and their nymphs in the following year. The general pattern holds when using an apple production baseline of 1964 instead of 1997.

5.4 Cicada and Infant Mortality: Instrumental Variable

Next I use cicada-apple interaction as an instrument for insecticide use in order to assess the resulting impact on infant mortality. The objective of the instrument is to isolate variation in insecticide use driven by cicada emergence in order to address endogeneity or omitted variable concerns. For example, factors like farm income can affect both insecticide use and infant mortality.

A major trade-off in using this instrument is that the time series is limited from 1992 to 2016 when USGS pesticide data are available. In addition to concerns about amplifying measurement error in the pesticide data, in many cases a county may only receive one 'treatment' from a 17-year cicada brood. This is in contrast to the main specification in Table 5 which uses a longer time series from 1950 to 2016 to test the impact of cicadas on infant mortality, as determined by tree crop intensity rather than insecticide use.

The IV results are contained in Table 6. For reference, OLS estimates are shown in the first two columns, in which Model (2) replicates Table 4's Model (1). Models (3)-(4) show

described in Bailey et al. 2016.

[9] As described earlier, I limit the leads and lags to four years to limit the distortion of the event study from the fact that many counties receive more than one cicada brood.

the IV results with instruments constructed using a dummy for top decile of apple production, and Models (5)-(6) use apple production in bushels. There is a consistent positive relationship between insecticides on infant health. All coefficients are larger and more precise than the OLS estimates, but the coefficients tend to vary and can become less precise depending on the specification, sample restrictions, and the fixed effects included.

The IV analysis indicates that each additional unit of insecticide can increase next-year infant mortality by 0.01 to 0.015. The standard deviation in insecticide use over time, after demeaning at the county level, is 24 kg per km^2. So a one standard deviation increase in insecticide use would increase infant mortality by 0.24 to 0.36. For context, this is a 4-6% increase over the current average infant mortality rate in the US of six deaths per thousand.

5.5 Interpretation of Infant Mortality Impact

Infant mortality has decreased by 80% over the course of this study, from a national average of 29 deaths per thousand to the current average of six, so the interpretation of coefficient magnitudes depends on the span of the sample. For the period of the instrumental variable analysis with the limited sample from 1992 to 2016, the average infant mortality is seven. For the longer timeframe from 1950 to 2016, the average is 15. This warrants some caution when interpreting and comparing coefficient magnitudes.

Table 1 shows that among top decile apple counties, insecticide use increases during a cicada emergence by 5 kg/km^2. Table 5 shows that these same treated counties see an increase in next-year infant mortality by 0.3 deaths per thousand, or a 2% increase over the full sample average infant mortality rate of 15. In terms of insecticide use, each additional kilogram of insecticide can be equated, therefore, to an increase in the infant mortality rate by 0.06, or 0.4%.

This result is about double the magnitude of the IV result, where each additional kilogram of insecticide increased infant mortality by 0.0125, on average, or 0.2% of the average infant mortality rate of seven deaths per thousand during 1992-2016. For context, mean insecticide use across counties and over time is 10 kg/km^2, so one more kilogram represents an approximate 10% increase over the sample mean.

Linking these results to a specific type of insecticide is challenging because I use an aggregate measure of insecticide use that sums up 160 insecticide constituents by weight. Further, there is little evidence that orchard growers and farm managers consistently choose one type of insecticide over others for cicada control, especially given that pest management practices vary greatly across the US and over time.

5.6 Timing and Sub-annual Impacts

Figure 7 shows the impact on infant mortality by quarter. This analysis is further limited to the period from 1995 to 2016, when Linked Infant Birth/Death Files are available that allow for sub-annual aggregation. There is an overall positive next-year impact that aligns with the main results of the longer-duration analyses from 1950-2016 in Table 5 and Figure 6, implying that the cicada-insecticide relationship to infant mortality holds in later time periods. The effect is largest in the third quarter of the year that follows, which is one quarter later than the results in Figure 5, which simply regresses quarterly infant mortality on insecticide use—notwithstanding cicadas and land use.

One explanation for this difference is that insecticide practices differ between row crops and tree crops. The average national effect would be dominated by row crops since apples account for only 1.4% of pesticide use in the US, while crops like corn, soy, cotton, potatoes, sorghum and wheat account for 86% (Fernandez-Cornejo et al. 2014). Insecticides are generally applied to row crops in the spring and summer and would affect early stage-pregnancies and infants born earlier in the following year. For orchards experiencing a cicada emergence, cicadas arrive in the summer and much of the insecticide application also targets the nymphs, which cause significant the damage (Hamilton 1961, Lloyd and White 1987). In this case, pesticide application would likely continue into the fall and through the next spring and summer, thus affecting birth outcomes later in the next year as well as the year afterwards.

5.7 Other Infant Health Outcomes

Next I assess infant health impacts beyond infant mortality. Using NCHS Natality Data files from 1968 to 2016, I compute three binary measures of infant health. The first is Apgar score (dummy for a score below 7 out of 10), a measure of the health of newborns based on a quick assessment of infant appearance, pulse, grimace, activity, respiration (hence acronym, Apgar). The second is premature birth (dummy if gestation period is under 37 weeks, the cutoff for premature birth). The last is birthweight (dummy if under 2500 grams, the threshold for low birthweight).

Table 7 shows that the cicada-apple interaction has a small but positive impact on the probability of a low Apgar score and premature birth, and a less clear relationship with low birthweight. All interaction coefficients are positive. Overall, these results are consistent with a story of insecticide exposure increasing the probability of infant death in the first year.

5.8 Education and Long-Term Impacts

I now look at the potential impact on educational achievement via elementary school cohorts exposed to a cicada emergence during conception or during the first year of life. Table 8 shows the impact on county-level scores in math and English language arts using Stanford Education Data Archives NAEP-equivalent test scores (Reardon et al. 2018). I pool county scores by cicada exposure cohorts, i.e., averaging the scores of third graders 9 years after a cicada event, fourth graders 10 years after, and fifth graders 11 years after. Models (2) and (4) include leads and lags of the cicada dummy to test the temporal relationship. While the leads and lags add some noise, there is a clear decline in average test scores of 1 to 1.3 NAEP-equivalent points among exposed cohorts. Each successive grade level NAEP score is, on average, 10 points higher, so this coefficient can be interpreted as a reduction of 10-13% of one grade-level's worth of learning.

Next I look at even longer-term impacts: whether cohorts conceived during a cicada emergence in tree crop-intensive counties experience a change in educational attainment. Using NCES data, I calculate the average dropout rate across school districts at a county-year level from 1991 to 2009. Table 9 shows the results of regressing the twelfth grade dropout rate on the cicada-apple interaction term. Model (2) includes long-term cicada lags ranging from 16 years after emergence to 22 years. There is a significant increase in the dropout rate 19 years later of exposed cohorts conceived during a cicada exposure, which is when these students would most likely be in the twelfth grade. The lag terms from 16 to 18 years are also positive but of a smaller magnitude, implying that there may be effects on exposed infants and toddlers.

The median twelfth grade dropout rate during this period is four per hundred students, and the standard deviation in apple bushel production in 1997 is 0.23 thousand bushels/km^2. Therefore, in the event of a cicada emergence, counties with one standard deviation higher apple intensity see an increase in the future dropout rate by 0.19, or about 5%. The same results, however, are not found when using a dummy for top apple production decile instead of production intensity.

It is important to note that the composition of counties over time is unknown. Since many people move in and out of counties over the course of two decades, it is not possible to know if those conceived during a cicada emergence were the same individuals in the county taking the elementary school tests and attending high school. Nevertheless, these findings align with Rauh et al. 2012 and provide evidence that insecticides can have long-term cognitive impacts that affect life outcomes beyond just infant health.

6 Water Mechanism

It may be surprising that fruit tree acreage, given its small footprint, can produce effects that are measurable at the county level. The largest apple producer in our sample, Wayne County, NY, has about 20,000 acres of apples trees, which is less than 5% of its land area. This is a small fraction compared to counties that intensively grow soy and corn, where row crops comprise a majority of the land. As such, only a small proportion of any county's population would be in close physical proximity to orchards.

Aside from farm workers directly exposed to insecticides, the primary channel in which a population is exposed to insecticides is likely water.[10] Insecticides are known to run off from agricultural fields into streams and groundwater. USGS found pesticides present in 54% of the 1,034 shallow groundwater sites sampled from 1993 to 1995 across 20 major hydrologic basins in the US (USGS 2019).

Private water well users, as opposed to those on centralized public systems, would be more exposed to insecticides via this water channel. USGS reports that 15% percent of the US population gets its drinking water from private wells, where the quality and safety are not regulated, yet 26% of people in the average US county get their water from private wells. This county-level figure is higher than the national average because people in densely populated areas tend to be supplied by public systems, while most private well users are in rural areas.

Using USGS data, I test whether there is heterogeneity in birth outcomes based on wa-

[10] Insecticides can also spread through aerial drift with wind.

ter source. The measure of water exposure is the proportion of the population in a county that is reliant on private wells. I take the average from 1985 to 2015 and create an indicator based on whether the county is below the national median ('LOW') or above ('HIGH'). Table 10 shows the results. Models (1) and (3) replicate the primary specification in Table 5 for comparison. Models (1)-(2) use a dummy for top-decile apple counties, while Models (3)-(4) use level of apple production. Apple-intensive counties with more private wells tend to have larger and more precise infant mortality coefficients than those with lower well use.

7 Robustness Checks

7.1 Yield Impacts

There are certain factors that could undermine the cicada-infant mortality story. Plausible candidates need to affect tree crop-intensive counties in the year following cicada emergence in ways that are different than those same counties in other years, as well as other tree crop-intensive counties in that same year that did not experience a cicada emergence.

One candidate is agricultural yields. If cicadas decimate apple production, for example, there could be a health impact via an economic channel. Our main dataset comes from the agricultural census, which occurs approximately every 5 years, and thus does not allow for testing annual shocks. USDA does, however, track annual apple production for a subset of 170 counties in the states of Virginia, South Carolina, Kansas, Pennsylvania, and New Jersey from 1972 to 2012. Using this limited data, I regress county-level apple production on leads and lags of cicada emergence.

Figure 8 plots the coefficients, with level of production on the top panel and log production on the bottom panel. While there is no significant relationship with level of production, the log values show a decrease in apple production in the year before and the year of cicada emergence. A weaker but non-significant effect seems to persist afterward. This is unsurprising and partly justifies why orchard owners apply insecticides. It also aligns with the agronomic and ecological literature showing that cicadas reduce tree growth, with feeding nymphs being a major main culprit (Karban 1982).[11] This negative yield impact, however, is less than the 30%-plus reduction in tree growth observed in natural settings in the absence of insecticides.

There are two main reasons that the economic channel is unlikely to explain the infant mortality relationship. First, yields decline in the year before and the year of a cicada emergence, but there is no evidence of an increase in infant mortality until the year afterwards. Second, tree crops comprise a very small portion of the economic value of most counties. For example, Wayne County, NY, the largest apple producer in the eastern half of the US, has a county GDP in 2012 of $3 billion according to the Bureau of Labor Statistics. The combined value of all fruit production is just $79 million according to USDA NASS, just 2.5% of GDP. Taken together, it seems unlikely that a yield-based economic channel is the main driver of observed health impacts, especially ones that are averaged over an entire county.

7.2 Births

The other factor that could complicate the cicada-infant mortality story is if cicada emergence alters behavior in ways that affect birth outcomes outside of the insecticide channel (e.g., if cicadas make people engage in more or less risky behavior). Since cicada emergences are short-lived, generally lasting only four to five weeks, it seems unlikely that cicadas would *in themselves* alter average outcomes at the county level over the course of

[11] Nymphs feed strongly on roots both leading up to emergence and afterwards in the first years of establishment

the entire following year. Further, one would have to believe that people in counties with a high proportion of tree crops behave differently in response to cicadas than people in counties with fewer tree crops—a proposition that seems unlikely.

Table 11 presents the results of a regression of next-year birth rate on cicada emergence and apple intensity. Birth rate is computed with ICPSR natality data as total annual births per thousand people (crude) and thousand women of child-bearing age (ages 15-44). The interaction coefficients in Models (2) and (4) are close to zero and insignificant. Behavior, as it relates to number of births, is not different in apple-intensive 'treated' counties versus untreated counties, thus not undermining the insecticide health relationship.

However, a relationship seems to exist where births increase the year after cicada emergence. This interesting finding holds after controlling for fixed effects at the county, year, and state-year level. I calculate a back-of-the-envelope using the crude birth rate estimate of 0.11 per thousand and the fact that the 1950 to 2016 population of counties with a cicada presence in my sample averaged 87 million. Since cicadas emerge every 16 years on average (3 broods have 13-year cycles, 12 broods have 17-year cycles), this means that an additional 600 people are born each year in the US, on average, because of cicadas.

This strange result could reflect something similar to the dynamic found in Evans et al. 2010 and Burlando 2014 where birth rates increase after hurricanes (when people are forced to stay inside) or power outages. Or perhaps there is a physiological effect that science has yet to uncover, one that occurs when humans witness millions of frenzied creatures emerging from over a decade underground only to live for a few weeks, just long enough to sing a shrill song, mate, and die.

7.3 Maternal characteristics

Finally, there may be concerns that the composition of mothers may somehow change. In other words, maybe the mothers in fruit-tree intensive counties who give birth in the year following cicada emergence are somehow different in ways that could explain some of the health outcomes. To control for this, I rerun the main specification from Equation 2 with several maternal controls, including age, education, race, weight gain during pregnancy, and cigarette consumption during pregnancy in Table 12. This limits the time series from 1990 to 2009 when all these variables were tracked on birth certificates. Model (1) replicates the main specification from Model (4) in Table 5 for reference. Model (2) performs the same analysis over the restricted sample. Models (3)-(5) show that the cicada-apple interaction coefficient remains significant and little changed with maternal controls.

8 Conclusions

Insecticides are critical to agricultural productivity, but they also pose risks to the population that are difficult to measure. In this paper, I use the mass emergence of periodical cicadas in 13 and 17-year cycles to identify the impact of insecticides on human health.

I find an increase in insecticide use in counties experiencing a cicada emergence that is limited to areas with a large amount of woody crops (i.e., fruit trees), as opposed to herbaceous row crops like corn and soy. This is because cicadas only damage woody plants: nymphs feed on tree roots, and adult cicadas lay their eggs in small branches.

I exploit this variation to compare treated counties (i.e., counties with high levels of apple production that experience a cicada emergence) to untreated counties. In the treated counties, we see a jump in next-year infant mortality by 0.3 deaths per thousand births. The birth impact extends into the second year as farmers continue applying higher levels of insecticide to control cicada nymph establishment. Sub-annual impacts align with the timing and patterns of insecticide usage by farmers.

Treated counties also have lower infant Apgar scores and more premature births. There is also evidence of long-term cohort effects in the form of lower elementary school test scores and higher high school dropout rates. I then demonstrate that insecticide exposure is likely through the water channel.

The findings may be generalizable beyond just very agriculturally-intensive areas. Tree crops cover a relatively small portion of US counties and total pesticide use is relatively low in most tree-intensive counties, especially compared to row crop agriculture. Together this supports the idea that moderate levels of pesticides, not just extreme exposures, can have negative health impacts.

Overall this paper contributes to the environmental and health economics literature on the health impacts of agricultural inputs. While acknowledging the large benefits of pesticides to agricultural productivity, the findings warrant caution in the over-application of insecticides. This paper also provides an example of how ecological phenomena like cicadas may be used to generate quasi-random variation that can be employed to answer important economic and public health questions.

References

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–172. ISSN: 0895-3309, accessed May 31, 2019. doi:10.1257/jep.25.3.153. https://www.aeaweb.org/articles?id=10.1257/jep.25.3.153.

Andersson, Henrik, Damian Tago, and Nicolas Treich. 2014. *Pesticides and health: A review of evidence on health effects, valuation of risks, and benefit cost analysis.* TSE Working Paper 14-477. Toulouse School of Economics (TSE). Accessed September 12, 2019. https://econpapers.repec.org/paper/tsewpaper/27991.htm.

Bailey, Martha, Karen Clay, Price Fishback, Michael R. Haines, Shawn Kantor, Edson Severnini, and Anna Wentz. 2016. *U.S. County-Level Natality and Mortality Data, January 1915-December 2007: Version 2.* https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/36603/versions/V2.

Bell, Erin M., Irva Hertz-Picciotto, and James J. Beaumont. 2001. "A Case-Control Study of Pesticides and Fetal Death Due to Congenital Anomalies." *Epidemiology* 12 (2): 148. ISSN: 1044-3983, accessed September 12, 2019. https://journals.lww.com/epidem/fulltext/2001/03000/a_case_control_study_of_pesticides_and_fetal_death.5.aspx?casa_token=fcNybbk_PHsAAAAA:6lsK9JWxd0xLPr7IP3ig7kDBlgtuCYURTLhxbtPvN5alyLm3Q3SFK0skTyez5b1Bi65lI3mNTglP9GFAxUCc_9QK.

Brainerd, Elizabeth, and Nidhiya Menon. 2014. "Seasonal effects of water quality: The hidden costs of the Green Revolution to infant and child health in India." *Journal of Development Economics* 107 (C): 49–64. Accessed September 12, 2019. https://ideas.repec.org/a/eee/deveco/v107y2014icp49-64.html.

Buffington, E.J., and S.K. Mcdonald. 2006. *Banned and Severely Restricted Pesticides, CEPEP, Colorado State University.* https://webdoc.agsci.colostate.edu/cepep/FactSheets/141BannedPesticides.pdf.

Burlando, Alfredo. 2014. "Power Outages, Power Externalities, and Baby Booms." *Demography* 51 (4): 1477–1500. ISSN: 1533-7790, accessed July 1, 2019. doi:10.1007/s13524-014-0316-7. https://doi.org/10.1007/s13524-014-0316-7.

Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics* 118 (3): 1121–1167. https://academic.oup.com/qje/article/118/3/1121/1942999.

Dybas, Henry S., and D. Dwight Davis. 1962. "A Population Census of Seventeen Year Periodical Cicadas." *Ecology* 43 (3): 432–444. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1933372.

Evans, Richard W., Yingyao Hu, and Zhong Zhao. 2010. "The fertility effect of catastrophe: U.S. hurricane births." *Journal of Population Economics* 23 (1): 1–36. ISSN: 1432-1475, accessed July 1, 2019. doi:10.1007/s00148-008-0219-2. https://doi.org/10.1007/s00148-008-0219-2.

Fernandez-Cornejo, Jorge, Richard F. Nehring, Craig Osteen, Seth Wechsler, Andrew Martin, and Alex Vialou. 2014. "Pesticide Use in U.S. Agriculture: 21 Selected Crops, 1960-2008." *SSRN Electronic Journal.* http://www.ssrn.com/abstract=2502986.

Frank, Eyal. 2018. "The Effects of Bat Population Losses on Infant Mortality through Pesticide Use in the U.S." *Unpublished Working Paper.*

Garry, Vincent F, Mary E Harkins, Leanna L Erickson, Leslie K Long-Simpson, Seth E Holland, and Barbara L Burroughs. 2002. "Birth defects, season of conception, and sex of children born to pesticide applicators living in the Red River Valley of Minnesota, USA." *Environmental Health Perspectives* 110 (Suppl 3): 441–449. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1241196/.

Haines, Michael, Price Fishback, and Paul Rhode. 2014. *United States Agriculture Data, 1840 - 2012: Version 4.* https://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/35206/versions/V4.

Hamilton, D. W. 1961. "Periodical Cicadas, Magicicada Spp., as Pests in Apple Orchards." *Proceedings of the Indiana Academy of Science* 71:116–121. http://journals.iupui.edu/index.php/ias/article/view/5930.

Jefferson, Thomas. 1944. *Thomas Jefferson's Garden Book, 1766-1824: With Relevant Extracts from His Other Writings.* American Philosophical Society.

Jurewicz, Joanna, Wojciech Hanke, Carolina Johansson, Christofer Lundqvist, Sandra Ceccatelli, Peter Van Den Hazel, Margaret Saunders, and Rolf Zetterstrom. 2006. "Adverse health effects of children's exposure to pesticides: What do we really know and what can be done about it." *Acta Paediatrica* 95 (s453): 71–80. ISSN: 1651-2227, accessed September 12, 2019. doi:10.1080/08035320600886489. https://onlinelibrary.wiley.com/doi/abs/10.1080/08035320600886489.

Karban, Richard. 1980. "Periodical cicada nymphs impose periodical oak tree wood accumulation." *Nature* 287 (5780): 326–327. https://www.nature.com/articles/287326a0.

———. 1982. "Experimental removal of 17-year cicada nymphs and growth of host apple trees." *Journal - New York Entomological Society (USA).* http://agris.fao.org/agris-search/search.do?recordID=US8230116.

Koenig, Walter D., Leslie Ries, V. Beth K. Olsen, and Andrew M. Liebhold. 2011. "Avian predators are less abundant during periodical cicada emergences, but why?" *Ecology* 92 (3): 784–790. http://doi.wiley.com/10.1890/10-1583.1.

Krawczyk, Grzegorz. 2017. "Tree Fruit Insect Pest - Periodical Cicada." *Penn State Extension.* https://extension.psu.edu/tree-fruit-insect-pest-periodical-cicada.

Larsen, Ashley E., Steven D. Gaines, and Olivier Deschenes. 2017. "Agricultural pesticide use and adverse birth outcomes in the San Joaquin Valley of California." *Nature Communications* 8 (1): 302. https://www.nature.com/articles/s41467-017-00349-2.

Lloyd, Monte, and JoAnn White. 1987. "Xylem Feeding by Periodical Cicada Nymphs on Pine and Grass Roots, With Novel Suggestions for Pest Control in Conifer Plantations and Orchards." 87:5.

NCHS. 2019. *National Vital Statistics System.* https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm.

Rauh, Virginia A., Frederica P. Perera, Megan K. Horton, Robin M. Whyatt, Ravi Bansal, Xuejun Hao, Jun Liu, Dana Boyd Barr, Theodore A. Slotkin, and Bradley S. Peterson. 2012. “Brain anomalies in children exposed prenatally to a common organophosphate pesticide.” *Proceedings of the National Academy of Sciences* 109 (20): 7871–7876. ISSN: 0027-8424, 1091-6490, accessed September 12, 2019. doi:10.1073/pnas.1203396109. https://www.pnas.org/content/109/20/7871.

Reardon, Sean F., Andrew D. Ho, Erin M. Fahle, Demetra Kalogrides, and Richard DiSalvo. 2018. *Stanford Education Data Archive (SEDA).* Accessed June 28, 2019. https://purl.stanford.edu/db586ns4974.

Regidor, E., E. Ronda, A. M. Garcfffdfffda, and V. Domfffdfffdnguez. 2004. “Paternal exposure to agricultural pesticides and cause specific fetal death.” *Occupational and Environmental Medicine* 61 (4): 334–339.

Schreinemachers, Dina M. 2003. “Birth malformations and other adverse perinatal outcomes in four U.S. Wheat-producing states.” *Environmental Health Perspectives* 111 (9): 1259–1264. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1241584/.

US EPA, OCSPP. 2017. *Pesticides Industry Sales and Usage 2008 - 2012 Market Estimates.* Reports and Assessments. Accessed June 20, 2019. https://www.epa.gov/pesticides/pesticides-industry-sales-and-usage-2008-2012-market-estimates.

USGS. 2019. *NAWQA The Pesticide National Synthesis Project.* https://water.usgs.gov/nawqa/pnsp/usage/maps/county-level/.

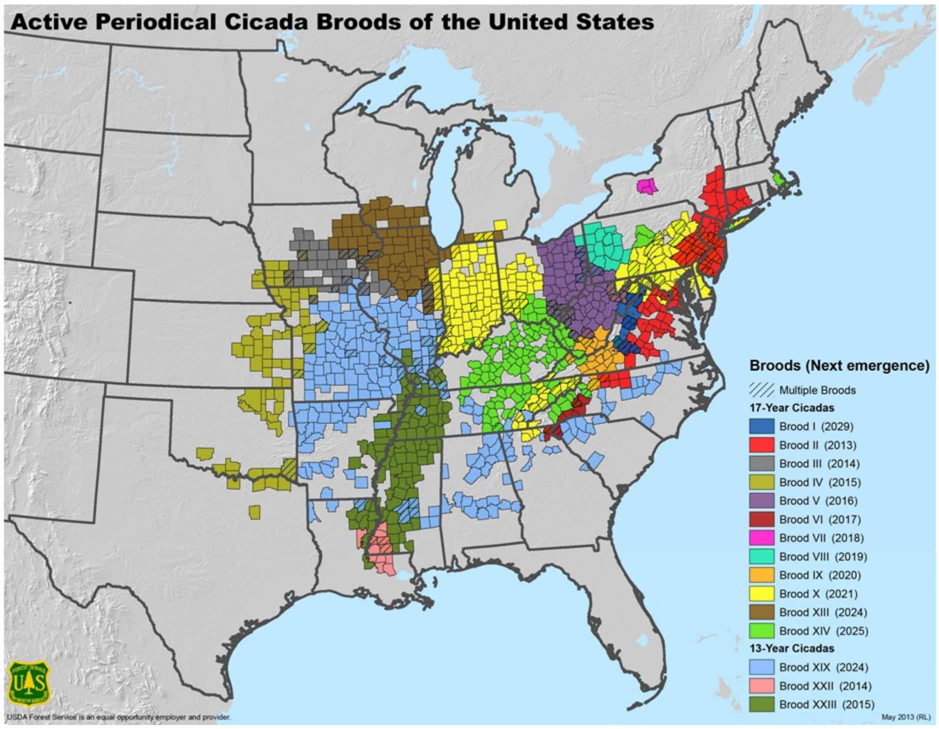
Williams, K S, and C Simon. 1995. “The Ecology, Behavior, and Evolution of Periodical Cicadas”: 29.

Winchester, Paul D, Jordan Huskins, and Jun Ying. 2009. “Agrichemicals in surface water and birth defects in the United States.” *Acta Paediatrica (Oslo, Norway : 1992)* 98 (4): 664–669. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2667895/.

Zheng, T, J Zhang, KE Sommer, BA Bassig, XC Zhang, J Braun, SQ Xu, et al. 2016. “Effects of environmental exposures on fetal and childhood growth trajectories.” *Annals of global health* 82 (1): 41–99. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5967632/.

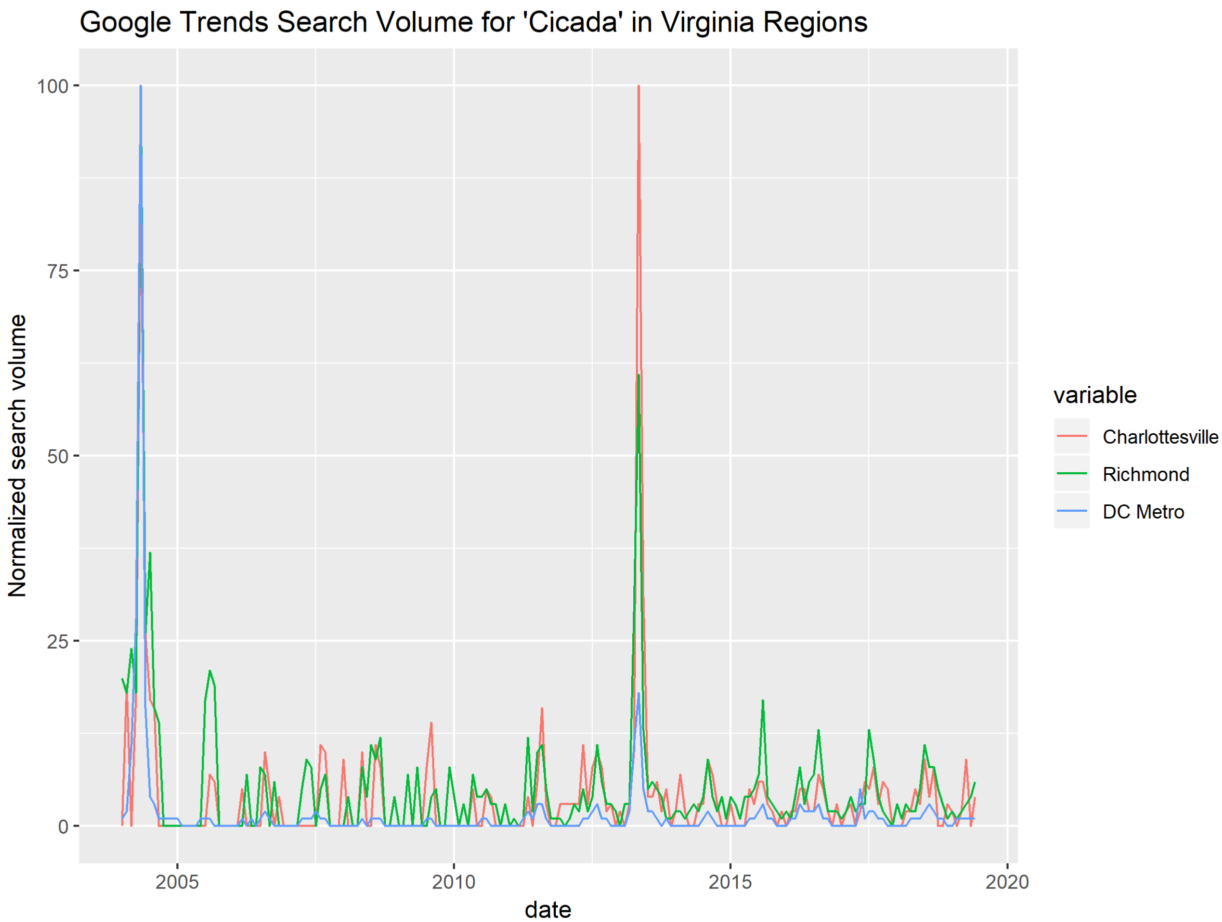
Figures

Figure 1



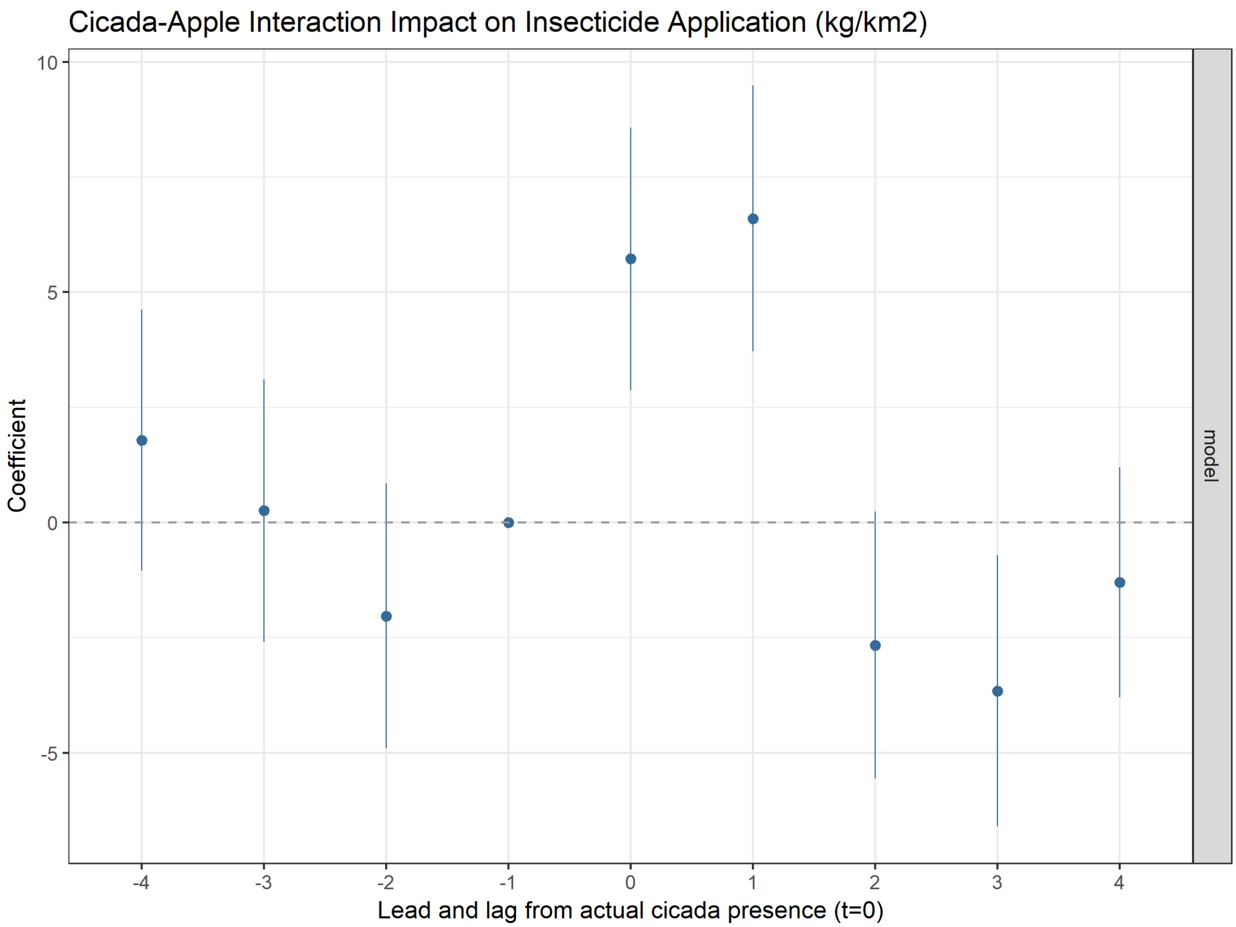
Source: Liebhold, A. M., Bohne, M. J., and R. L. Lilja. 2013. USDA Forest Service Northern Research Station.

Figure 2



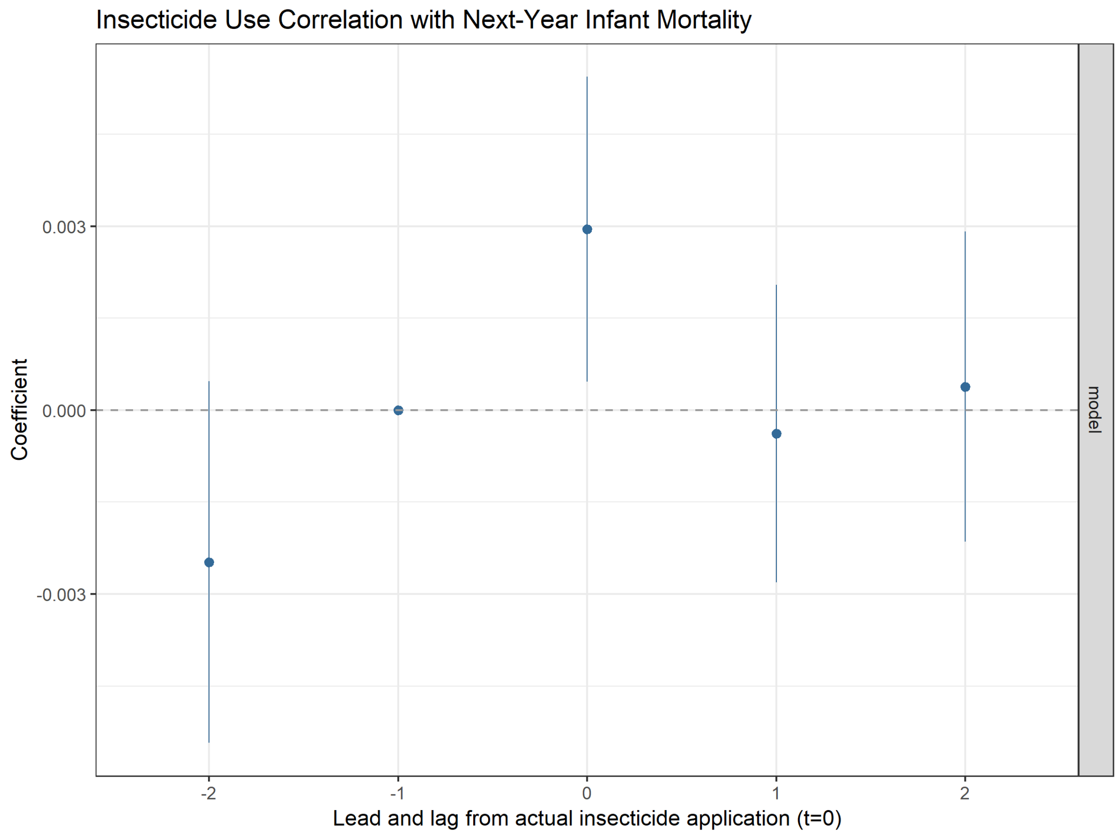
Source: Google Trends

Figure 3



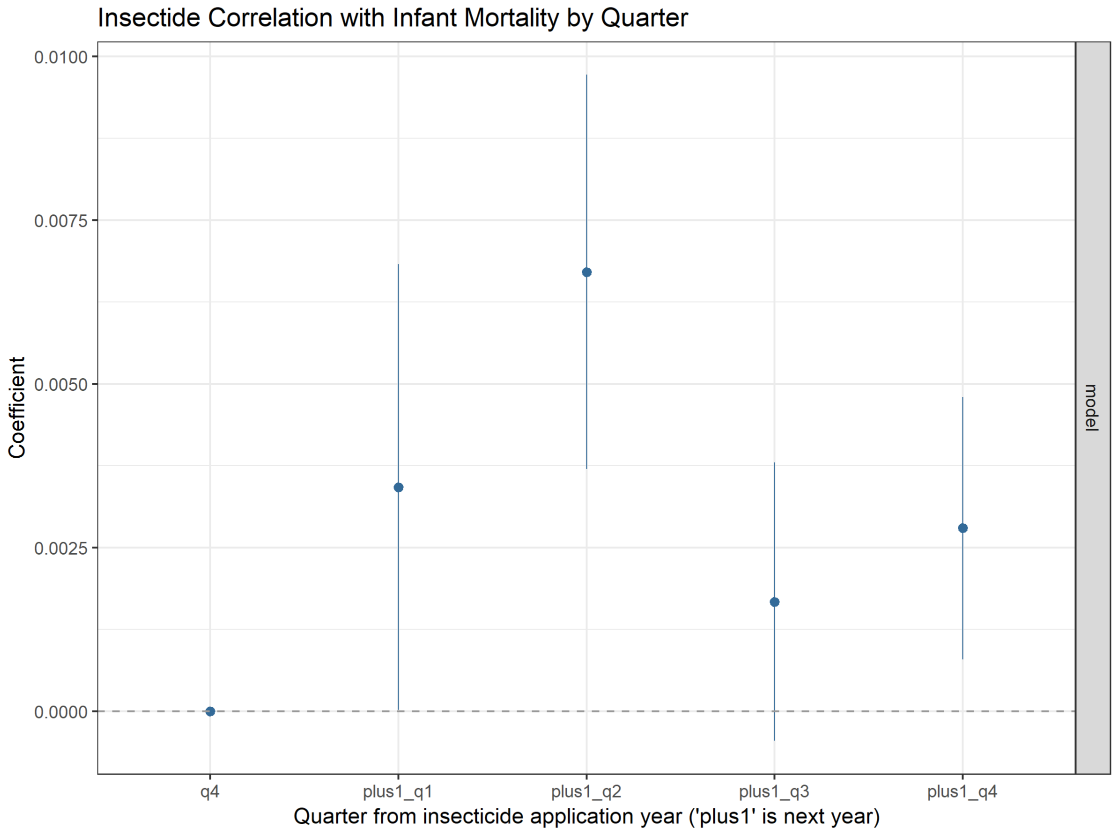
Notes: Event study based on Model (4) from Table 1 but includes cicada leads and lags. Time series limited to availability of USGS pesticide data, 1992 to 2016. Observations with no leading or lagging cicada events during the sample time period are excluded to balance the panel. Solid lines show 95% confidence intervals. Normalized to the year before cicada emergence.

Figure 4



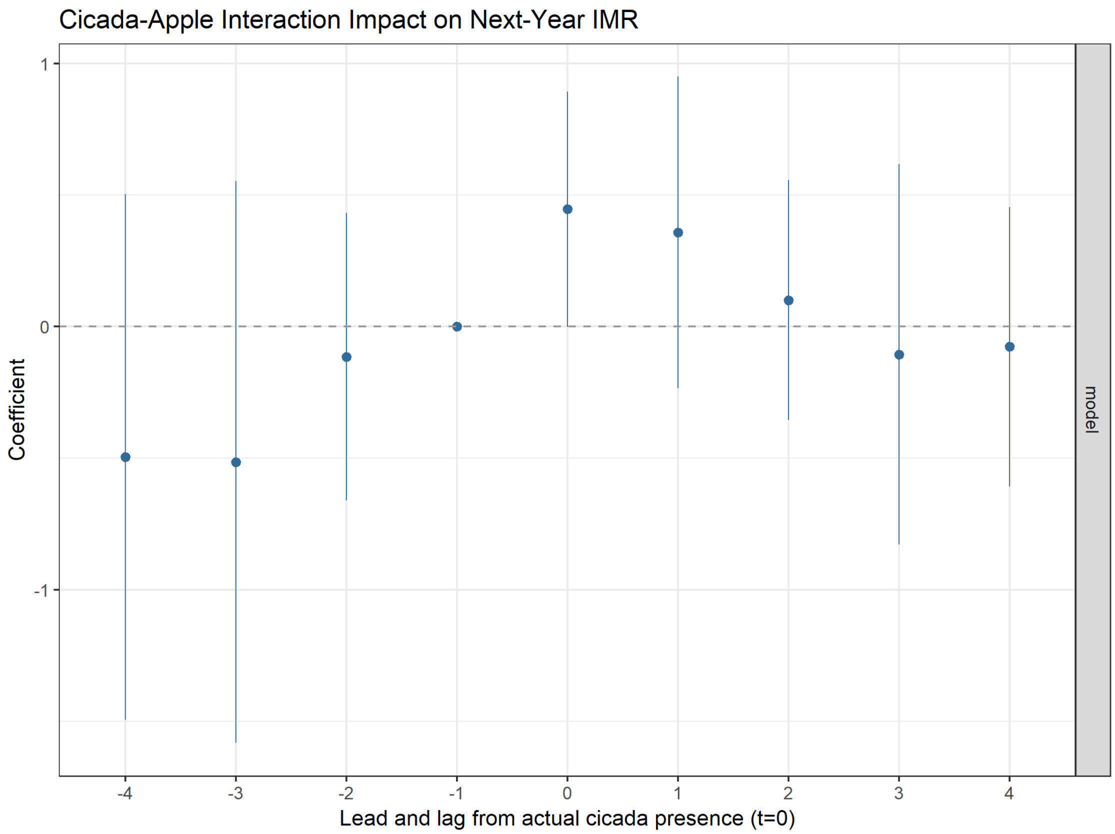
Notes: Event study based on Model (3) from Table 4 with the inclusion of two leads and lags of insecticide use. Time series limited to availability of USGS pesticide data, 1992 to 2016. Excludes observations with zero infant deaths. Solid lines show 95% confidence intervals. Normalized to the year before current year insecticide use.

Figure 5



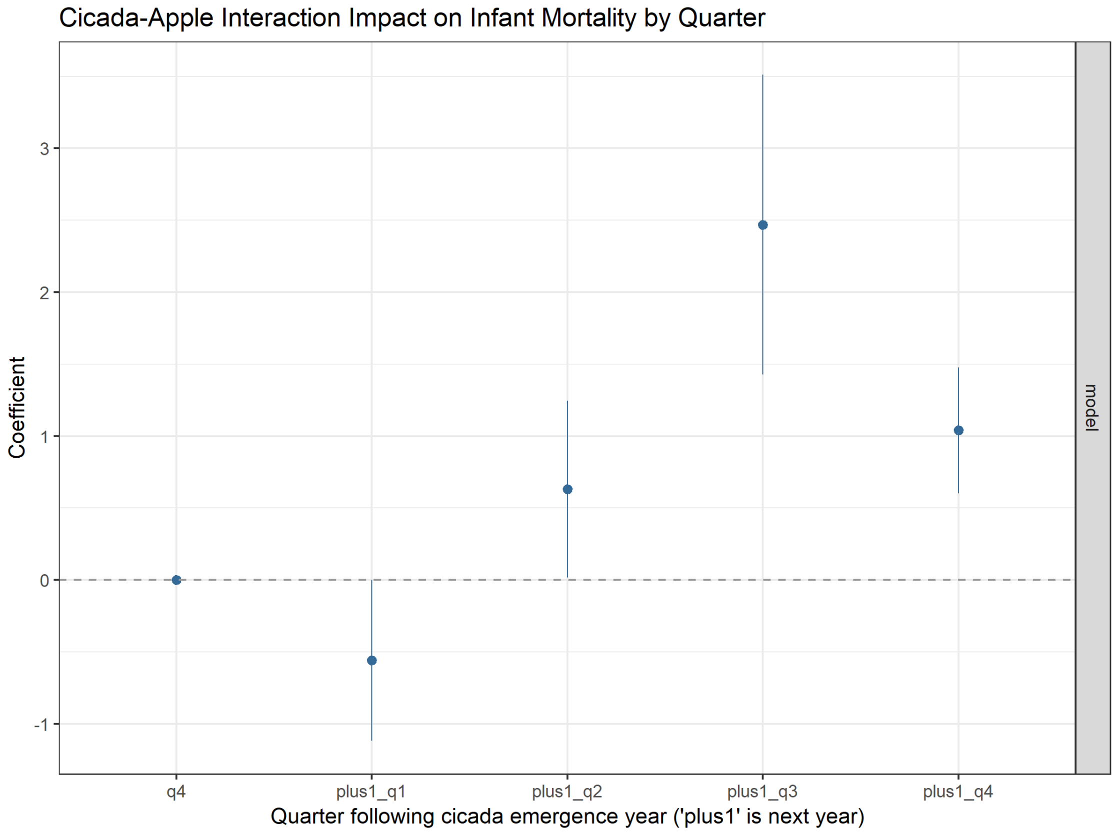
Notes: Event study based on models in Table 4 but run at the quarterly level. Time series limited to availability of USGS pesticide and sub-annual Infant Birth/Death data, 1995 to 2016. Excludes observations with zero infant deaths. Quarterly fixed effects included to address naturally-occurring fixed differences by season. Solid lines show 95% confidence intervals. Normalized to the fourth quarter of insecticide application year.

Figure 6



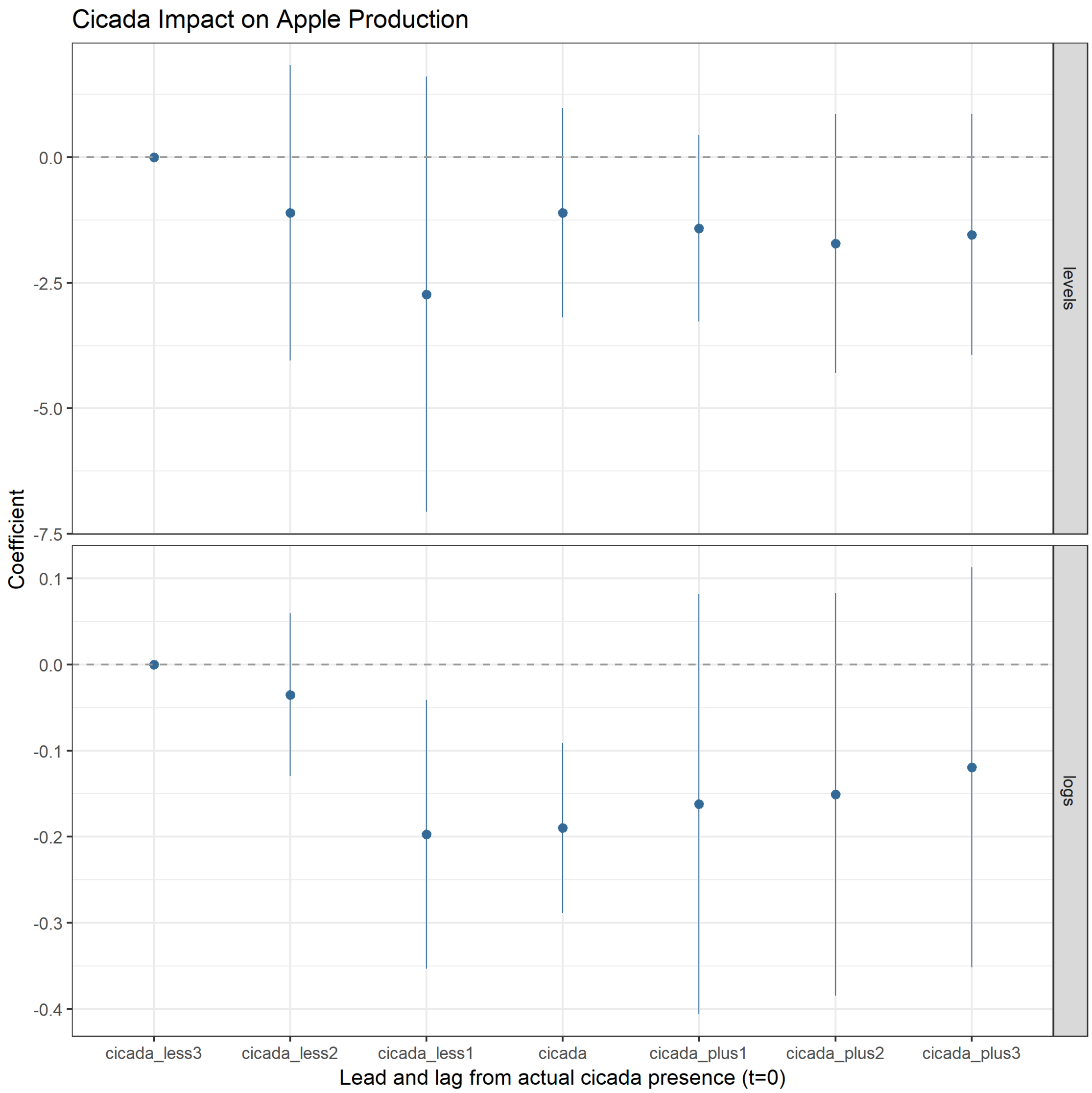
Notes: Event study based on Model (4) of Table 5, but including cicada leads and lags. Observations with no leading or lagging cicada events during the sample time period are excluded to balance the panel. Solid lines show 95% confidence intervals. Normalized to the year before cicada emergence.

Figure 7



Notes: Event study is based on models in Table 5 but run at the quarterly level. Time series limited to availability of sub-annual Infant Birth/Death data, 1995 to 2016. Excludes observations with zero infant deaths. Quarterly fixed effects included to address naturally-occurring fixed differences by season. Solid lines show 95% confidence intervals. Normalized to the fourth quarter of the cicada emergence year.

Figure 8



Notes: Event study of cicada impact on apple production. Dependent variable is county-level apple production in millions of bushels. Upper panel is levels, lower panel is log values. Annual time series is from 1972 to 2011 for select states with annual production data. County and state-year level fixed effect dummies included. Solid lines show 95% confidence intervals. Normalized to three years before cicada emergence.

Tables

Table 1: Cicadas and Insecticides

	Dependent variable:				
	Insecticide use (kg/km2)				
	(1)	(2)	(3)	(4)	(5)
cicada	0.08	−1.05	−0.82	−0.66	−0.86
	(0.70)	(0.73)	(0.74)	(0.74)	(0.80)
cicada:fruit_acres_HIGH		9.78***			
		(1.90)			
cicada:apples_acres_HIGH			6.20***		
			(1.74)		
cicada:bushels_HIGH				4.89***	
				(1.71)	
cicada:factor(bushels_quantile)1					−0.71
					(2.20)
cicada:factor(bushels_quantile)2					3.30
					(2.20)
cicada:factor(bushels_quantile)3					3.62*
					(2.14)
cicada:factor(bushels_quantile)4					4.86**
					(2.11)
County FE	X	X	X	X	X
State-Year FE	X	X	X	X	X
Observations	60,192	60,192	60,192	60,139	60,192
R^2	0.46	0.46	0.46	0.46	0.46

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is county-level insecticide intensity defined as insecticide use divided by county land area. Insecticide use is the combined sum of the USGS EPest-high values for the 160 constituents with insecticidal properties. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Models (2)-(4) include dummies for the top decile counties in fruit tree intensity based on fruit production per land area in 1997 (between 242 to 246 treated counties in eastern half of the US). Model (5) divides counties with positive apple production into quantiles. Time series limited to availability of USGS pesticide data, 1992 to 2016. County and state-year level fixed effect dummies included in all models.

Table 2: Falsification by Pesticide Type (kg/km2)

	Dependent variable:			
	Insecticide		Herbicide	Fungicide
	(1)	(2)	(3)	(4)
cicada	−0.66	−0.65	0.67	−0.33
	(0.74)	(0.74)	(0.83)	(0.76)
cicada:bushels_HIGH	4.89***	4.88***	−2.68	−0.15
	(1.71)	(1.71)	(1.91)	(1.76)
Same sample		X	X	X
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	60,139	59,798	59,798	59,798
R^2	0.46	0.46	0.86	0.57

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is county-level pesticide use divided by county land area. Pesticide use is the combined sum of the USGS EPest-high values for constituents with insecticidal, herbicidal, and/or fungicidal properties. Many pesticides had multiple properties. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Apples bushels HIGH is a dummy for the top decile counties in apple production in 1997. Time series limited to availability of USGS pesticide data, 1992 to 2016. To ensure comparability, Models (2)-(4) restrict sample to county-year observations with complete insecticide, herbicide, and fungicide data. County and state-year level fixed effect dummies included in all models.

Table 3: Falsification by Crop (kg/km2)

	Dependent variable:			
	Insecticide use (kg/km2)			
	(1)	(2)	(3)	(4)
cicada	−0.66	−0.94	−0.05	0.09
	(0.74)	(0.81)	(0.79)	(0.79)
cicada:apples_HIGH	4.89***	5.27***		
	(1.71)	(1.86)		
cicada:corn_HIGH			−1.19	
			(2.32)	
cicada:soy_HIGH				−2.90
				(2.31)
Same sample		X	X	X
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	60,139	53,195	53,195	53,195
R^2	0.46	0.46	0.46	0.46

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is county-level insecticide intensity defined as insecticide use divided by county land area. Insecticide use is the combined sum of the USGS EPest-high values for the 160 constituents with insecticidal properties. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Apples HIGH is a dummy for the top decile counties in apple production in 1997. Corn HIGH and soy HIGH are top decile corn and soy counties, respectively. Time series limited to availability of USGS pesticide data, 1992 to 2016. To ensure comparability, Models (2)-(4) restrict sample to county-year observations with complete apple, corn, and soy data. County and state-year level fixed effect dummies included in all models.

Table 4: Insecticide Use Correlation with Infant Mortality

	Dependent variable:		
	Next-Year Infant Mortality Rate (IMR)		
	(1)	(2)	(3)
insecticide_use	0.0009	0.0008	0.0025*
	(0.0018)	(0.0017)	(0.0013)
Only 5-plus births		X	X
Only 1+ IMR			X
County FE	X	X	X
State-Year FE	X	X	X
Observations	53,584	45,769	37,508
R^2	0.1090	0.1455	0.5424

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Insecticide use is the combined sum of the USGS EPest-high values for the 160 constituents with insecticidal properties, divided by county area. Time series limited to availability of USGS pesticide data, 1992 to 2016. Models (2)-(3) exclude county-year observations with less than five births. Model (3) is further limited to observations with at least one infant death. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 5: Cicada Impact on Infant Mortality, 1950-2016

	Dependent variable:			
	Next-Year Infant Mortality Rate (IMR)			
	(1)	(2)	(3)	(4)
cicada	0.15	0.10	0.12	0.13
	(0.19)	(0.21)	(0.20)	(0.19)
cicada:bushels_HIGH		0.30*		
		(0.17)		
cicada:bushels_1964			0.58***	
			(0.18)	
cicada:bushels_1997				0.44**
				(0.18)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	143,976	141,814	141,694	141,814
R^2	0.53	0.53	0.53	0.53

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Excludes county-year observations with less than 5 births. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels HIGH is a dummy for the top decile counties in apple production in 1997, while bushels 1997 and bushels 1964 is the actual level of apple production in those years. Time series from 1950 to 2016. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 6: Cicada-Apple Interaction Impact Infant Mortality

	Dependent variable:					
	Next-Year Infant Mortality Rate (IMR)					
	OLS		IV - top decile		IV - production	
	(1)	(2)	(3)	(4)	(5)	(6)
insecticide_use	0.001	0.001	0.01^{**}	0.01	0.01^{*}	0.01^{*}
	(0.002)	(0.002)	(0.005)	(0.004)	(0.01)	(0.01)
5+ births		X		X		X
County FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
State-Year	X	X	X	X	X	X
Observations	53,584	45,769	53,533	45,721	53,480	45,672
R^2	0.11	0.15	0.10	0.13	0.10	0.13

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Models (2), (4), and (6) exclude county-year observations with less than 5 births. Models (1)-(2) show an OLS regression of infant mortality on insecticide use. Models (3)-(4) instrument insecticide use on the cicada-apple intensity interaction (dummy for top decile of apple production). Models (5)-(6) use county-level apple production as the instrument after dropping outlier observations above the 99.9th percentile of annual insecticide use. Time series limited to 1992 to 2016 when pesticide data available. County and state-year level fixed effect dummies included as specified, though state-year fixed effects are not separately included in the second stage. Standard errors clustered at state level.

Table 7: Cicada-Apple Interaction Impact on Other Birth Outcomes

	Next-year birth outcome		
	Prob. Low Apgar	Prob. Premature	Prob. Low Birthweight
	(1)	(2)	(3)
cicada	−0.0001	−0.0006	−0.0009
	(0.0008)	(0.0009)	(0.0007)
cicada:bushels	0.0015*	0.0024***	0.0013
	(0.0008)	(0.0007)	(0.0017)
County FE	X	X	X
State-Year FE	X	X	X
Observations	82,009	107,611	110,378
R^2	0.1206	0.2192	0.2730

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variables are various next-year birth outcomes averages at the county level: Apgar low is a dummy for a score below 7 out of 10 (time series from 1978 to 2016); Premature is a dummy if gestation is under 37 weeks (time series from 1968 to 2016); Birthweight low is a dummy if under 2500 grams (time series from 1968 to 2016). Excludes county-year observations with less than 5 births. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels is apple production in 1997. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 8: Cicada-Apple Interaction Impact on Elementary School Test Scores

	NAEP-equivalent average test scores			
	Math		English	
	(1)	(2)	(3)	(4)
cicada_less2:bushels_HIGH		−0.26		0.02
		(0.50)		(0.49)
cicada_less1:bushels_HIGH		0.07		0.01
		(0.48)		(0.46)
cicada:bushels_HIGH	−1.09*	−1.20*	−1.25**	−1.34*
	(0.57)	(0.65)	(0.58)	(0.68)
cicada_plus1:bushels_HIGH		−0.08		−0.05
		(0.39)		(0.58)
cicada_plus2:bushels_HIGH		−0.43		−0.76*
		(0.32)		(0.45)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	10,733	10,733	11,379	11,379
R^2	0.91	0.91	0.90	0.90

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is county-level averages of Stanford Education Data Archive's NAEP-equivalent test scores averaged for all elementary school students (grades 3-5) in the same 'cicada exposure cohort'. For example, scores are the average of 3rd graders 9 years after cicada exposure, 4th graders 10 years after cicada exposure, and 5th graders 11 years after cicada exposure. Annual scores available from Spring 2009 to Spring 2015. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels HIGH is a dummy for the top decile counties in apple production in 1997. Non-interacted cicada coefficients are excluded from output for brevity. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 9: Cicada-Apple Interaction Impact on Dropout Rates

	Dependent variable:	
	Dropout rate per 100 students	
	(1)	(2)
cicada_plus16:bushels		0.15**
		(0.07)
cicada_plus17:bushels		0.27*
		(0.15)
cicada_plus18:bushels		0.26
		(0.25)
cicada_plus19:bushels	0.77**	0.85***
	(0.32)	(0.31)
cicada_plus20:bushels		−0.13
		(0.41)
cicada_plus21:bushels		−0.003
		(0.16)
cicada_plus22:bushels		−0.10
		(0.15)
County FE	X	X
State-Year FE	X	X
Observations	22,716	22,716
R^2	0.23	0.23

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is 12th grade dropout rates. Dropout rates are averaged across school districts at a county-year level and available from NCES from 1991 to 2009. Cicada lags are a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year (i.e., cicada_plus19 denotes a cicada occurance 19 years before the year of the dropout observation). Bushels is the level of apple production in 1997. Coefficients of the non-interacted cicada lags are omitted from output for brevity. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 10: Cicada-Apple Interaction Impact on Infant Mortality via Water Channel

	Dependent variable:			
	Next Year Infant Mortality Rate (IMR)			
	(1)	(2)	(3)	(4)
cicada	0.10	0.10	0.12	0.13
	(0.21)	(0.21)	(0.20)	(0.20)
cicada:bushels_HIGH	0.30*			
	(0.17)			
cicada:bushels_HIGH:well_LOW		0.25		
		(0.25)		
cicada:bushels_HIGH:well_HIGH		0.33*		
		(0.19)		
cicada:bushels			0.58***	
			(0.18)	
cicada:bushels:well_LOW				−0.67
				(1.61)
cicada:bushels:well_HIGH				0.62***
				(0.15)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	141,814	141,814	141,694	141,694
R^2	0.53	0.53	0.53	0.53

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Excludes county-year observations with less than 5 births. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels HIGH is a dummy for the top decile counties in apple production in 1997. Bushels is the actual level of apple production in 1964. The LOW and HIGH values for proportion of the population using private wells describe whether the county is above or below the national county median for that indicator. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 11: Cicada-Apple Interaction Impact on Birth Rates

	Dependent variable:			
	All people (Crude)		Female Age-Specific	
	(1)	(2)	(3)	(4)
cicada	0.11***	0.12***	0.41**	0.38*
	(0.03)	(0.04)	(0.17)	(0.22)
cicada:bushels_HIGH		−0.05		0.17
		(0.05)		(0.53)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	139,990	139,990	139,990	139,990
R^2	0.84	0.84	0.74	0.74

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year birth rate. Models (1)-(2) show the crude birth rate (births per 1000 people). Models (3)-(4) show births per thousand women of child bearing age (ages 15-44). Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 12: Cicada-Apple Interaction Impact on Infant Mortality, Maternal Characteristics

	Dependent variable:				
	Next-Year Infant Mortality Rate (IMR)				
	(1)	(2)	(3)	(4)	(5)
cicada	0.13	−0.14	−0.13	−0.14	−0.13
	(0.19)	(0.23)	(0.23)	(0.23)	(0.23)
education			−0.01	−0.004	−0.01
			(0.07)	(0.07)	(0.08)
age			−0.04	−0.03	−0.02
			(0.07)	(0.07)	(0.07)
black				0.71	0.80
				(0.58)	(0.59)
cig					0.04
					(0.05)
wtgain					0.04
					(0.04)
cicada:bushels	0.44^{**}	0.48^{**}	0.49^{**}	0.49^{**}	0.49^{**}
	(0.18)	(0.23)	(0.22)	(0.22)	(0.23)
County FE	X	X	X	X	X
State-Year FE	X	X	X	X	X
Observations	141,814	33,020	33,020	33,020	33,020
R^2	0.53	0.16	0.16	0.16	0.16

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Excludes county-year observations with less than 5 births. Maternal characteristics are annual county averages for education, age, a dummy for race (black versus non-black), number of cigarettes smoked during first trimester, and maternal weight gain. Outliers were dropped above the sample 99.9th percentile value to control for erroneous entries. Cicada dummy takes the value of 1 if there is a cicada emergence in the county in that year. Bushels is county apple production in 1997. Model (1) shows the main model results from 1950 to 2016. Model (2) replicates this with the restricted sample from 1990 to 2009 with data that includes all maternal characteristics. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 13: Cicada Impact on Infant Mortality, 1950-2016, Weighted by Births

	Dependent variable:			
	Next-Year Infant Mortality Rate (IMR)			
	(1)	(2)	(3)	(4)
cicada	0.04	0.05	0.06	0.06
	(0.12)	(0.12)	(0.12)	(0.12)
cicada:bushels_HIGH		0.10		
		(0.12)		
cicada:bushels_1964			0.40*	
			(0.22)	
cicada:bushels_1997				0.37**
				(0.17)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	149,813	147,565	147,441	147,565
R^2	0.83	0.82	0.83	0.82

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Regression weighted by the number of county births. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels HIGH is a dummy for the top decile counties in apple production in 1997, while bushels 1997 and bushels 1964 is the actual level of apple production in those years. Time series from 1950 to 2016. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.

Table 14: Cicada Impact on Infant Mortality, 1950-1988

	Dependent variable:			
	Next-Year Infant Mortality Rate (IMR)			
	(1)	(2)	(3)	(4)
cicada	0.26	0.20	0.22	0.23
	(0.23)	(0.25)	(0.24)	(0.24)
cicada:bushels_HIGH		0.35*		
		(0.21)		
cicada:bushels_1964			0.64*	
			(0.38)	
cicada:bushels_1997				0.47*
				(0.25)
County FE	X	X	X	X
State-Year FE	X	X	X	X
Observations	95,766	94,445	94,359	94,445
R^2	0.44	0.44	0.44	0.44

Notes: Linear regression. Stars increasing for significance at 90th, 95th, and 99th confidence levels. Dependent variable is next-year infant mortality rate (deaths per 1000 live births). Excludes county-year observations with less than 5 births. Cicada is a dummy variable taking the value of 1 if there is a cicada emergence in the county in that year. Bushels HIGH is a dummy for the top decile counties in apple production in 1997, while bushels 1997 and bushels 1964 is the actual level of apple production in those years. Time series limited to 1950-1988, when infant mortality data is available for all counties. County and state-year level fixed effect dummies included in all models. Standard errors clustered at state level.