Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Health Economics 23 (2004) 1209–1236
JOURNAL OF HEALTH ECONOMICS
www.elsevier.com/locate/econbase

Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma

Matthew J. Neidell*

University of Chicago, CISES, 5734 S. Ellis Ave., Chicago, IL 60637, USA

Received 1 August 2003; received in revised form 1 April 2004; accepted 1 May 2004

Abstract

This paper estimates the effect of air pollution on child hospitalizations for asthma using naturally occurring seasonal variations in pollution within zip codes. Of the pollutants considered, carbon monoxide (CO) has a significant effect on asthma for children ages 1–18: if 1998 pollution levels were at their 1992 levels, there would be a 5–14% increase in asthma admissions. Also, households respond to information about pollution with avoidance behavior, suggesting it is important to account for these endogenous responses when measuring the effect of pollution on health. Finally, the effect of pollution is greater for children of lower socio-economic status (SES), indicating that pollution is one potential mechanism by which SES affects health.
© 2004 Elsevier B.V. All rights reserved.

JEL classification: I12; J13; J15; Q25

Keywords: Pollution; Health; Asthma; Socio-economic status; Avoidance behavior; Zip code fixed effect

1. Introduction

A primary objective of air quality policies around the world is to protect human health. However, many critics argue that air quality standards are set somewhat arbitrarily with inconclusive evidence of the specific health benefits and with inadequate considerations of the

* Tel.: +1 773 834 9945; fax: +1 773 834 1045.
E-mail address: mneidell@uchicago.edu.

0167-6296/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2004.05.002

costs to producers. Given that substantial costs to industry have been widely demonstrated,[1] in order to determine optimal policy intervention it is crucial to identify the associated benefits from improvements in air quality.

While many studies have focused on estimating a relationship between pollution and health, they have largely neglected to consider that pollution exposure is endogenously determined if individuals make choices to maximize their well being. People with high preferences for clean air may choose to live in areas with better air quality. People can respond to a wide range of readily available information on pollution levels by adjusting their exposure. Failing to appropriately account for such actions can yield misleading estimates of the causal effect of pollution on health.

This paper focuses on developing an empirical strategy for measuring the effect of pollution on health. Specifically, I look at the effect of air pollution on children's hospitalization for asthma. Childhood asthma is of particular interest for two reasons: (1) asthma is the leading chronic condition affecting children; and (2) current pollution standards are based on adult health responses to pollution and children face a greater risk from pollution exposure due to the sensitivity of their developing biological systems.

This study builds on earlier work in five ways. One, I develop a unique, monthly, zip code level data set by matching information about all individual hospitalizations in California between 1992 and 1998 to ambient pollution levels, meteorological data, and various demographic data. Two, I identify the effect of pollution using naturally occurring seasonal variations within zip codes. Since zip codes are a finely defined geographic area and the seasonal patterns in pollution are remarkably strong and diverse throughout California, this controls for many confounding factors that might affect asthma hospitalization rates. Three, I allow the effect of pollution to differ with the age of the child, as biological models suggest it might. Four, I collect data about public announcements of "smog alerts" in order to show empirically that it is important to account for the endogeneity of household responses to pollution. Five, to assess if the effect of pollution varies across different segments of the population, I allow the effect of pollution to differ with socio-economic status (SES), as measured by education levels in the zip code.

The primary finding of this paper is that of the pollutants considered, carbon monoxide (CO) has a significant effect on hospitalizations for asthma among children ages 1–18, while none of the pollutants considered has a clear impact on hospitalizations for infants. This discrepancy across age groups is possibly due to the complications inherent in diagnosing asthma in infants or differing degrees of avoidance behavior by age. The decline in pollution levels from 1992 to 1998 decreased asthma rates by between 5 and 14%, resulting in a savings of approximately $5.2 million in hospital expenses for asthma admissions in California in 1998 alone.

A second finding is that families display avoidance behavior by responding to smog alerts, indicating the importance of accounting for the endogeneity of family behavior when measuring the causal effect of pollution on health. The announcement of smog alerts decreases asthma hospitalizations by roughly 1%, while including these announcements raises the effect of O_3 on admissions, although O_3 does not appear to significantly affect

[1] See, for example, Greenstone (2002) for estimates on the costs of the Clean Air Acts on industrial activity in the United States.

admissions. This suggests that omitting avoidance behavior yields estimates that are a lower bound of the biological effect of pollution on health.

Third, not only are the coefficients measuring the effect of pollution larger for low SES children, but these children are also exposed to considerably higher levels of pollution. As a result, they suffer greater harm from pollution, although higher pollution levels explain only as much as 6% of the gap in admission rates for asthma. Although there are many remaining factors for explaining this gap, this suggests that pollution is one potential mechanism for the well-known relationship between SES and health—poorer families are unable to afford to live in cleaner areas, and their children's health suffers as a result.

The paper is laid out as follows. Section 2 provides some background information on asthma and its potential association with pollution. Section 3 discusses the conceptual framework and estimation strategy. Section 4 describes the data used for the analysis. Section 5 presents the econometric results. Section 6 concludes with a discussion.

2. Background

Approximately 5 million children in the United States have asthma. It is the leading specific reason for school absence and the most frequent cause of pediatric emergency room use and hospital admission (National Institute of Environmental Health Sciences, 1/2000). Asthma disproportionately attacks children of lower SES, and continues for most well beyond childhood (American Academy of Pediatrics, 2000). Most disconcerting is that reported asthma rates for children age 18 and younger have increased by more than 70% from 1982 to 1994 (American Academy of Pediatrics, 2000)[2].

Despite mounting public concern, the factors influencing this illness are not fully understood, especially for children. Medical research has demonstrated that asthma is both a chronic and acute illness. In the chronic aspect, an individual's airways are persistently inflamed and their immune system is hyper-responsive, but the causes of this remain largely unknown (American Lung Association, 2000). During an acute response, an irritant is inhaled that causes three changes to occur: muscular bands around the bronchioles constrict, the linings of the airway become inflamed, and excess mucus is produced. The irritants are believed to cause this because, by being recognized by the immune system as foreign, immunoglobin E (IgE), an antibody, is produced in response. IgE binds with mast cells—particular cells filled with chemical mediators, causing the release of some of the mediators in the mast cells (American Academy of Pediatrics, 2000). As a result of these changes in lung functioning, the airways are severely narrowed, and, by making it difficult to breathe, may lead to an asthma attack.

While there are many potential irritants[3], or asthma "triggers", outdoor air pollution has long been suspected a major culprit. Carbon monoxide is an odorless, colorless gas that bonds with hemoglobin more easily than oxygen, so that it reduces the body's ability to deliver oxygen to organs and tissues. Although the biological plausibility of a direct effect of CO on asthma is unlikely, because CO mainly comes from vehicle exhaust, a

[2] There is, however, much debate regarding this apparent rise in asthma. I discuss this is more detail below.

[3] These include include molds, pollens, animal dander, tobacco smoke, weather, and exercise.

likely explanation is that CO functions as a proxy for vehicle emissions (U.S. EPA, 2000). Nitrogen dioxide (NO_2) is a brown, reactive gas that irritates the lungs and may lower resistance to respiratory infections. Particulate matter (PM_{10}), which can take many forms, including ash and dust, is believed to cause the most damage from the smallest particles, since they are inhaled deep into the lungs (U.S. EPA, 2003). The mechanisms through which particles harm health are controversial, however the leading theory is that they cause an inflammatory response, which weakens the immune system (Seaton, 1995). Ozone (O_3), the major component of urban smog, is a highly reactive compound that damages tissue, reduces lung function, and sensitizes the lungs to other irritants. Motor vehicles are a major source of PM_{10}, NO_2, and especially of CO; as much as 90% of CO in cities comes from motor vehicle exhaust (U.S. EPA, 2000), while O_3 is formed through reactions between nitrogen oxides (such as NO_2) and volatile organic compounds (which are found in auto emissions, among other sources) in heat and sunlight.

Many researchers have attempted to estimate the link between these pollutants and childhood asthma, but with mixed results.[4] Most studies have been short time-series that focus on a given city and track the daily number of hospital or emergency room (ER) admissions for asthma and the average daily levels of various criteria pollutants.[5] A wide range of estimated correlations between admissions for asthma and CO, O_3, PM_{10}, and NO_2 have been reported, with no clear patterns or magnitude of effects evident.[6]

Due to the inconclusive findings and the fact that ambient air pollution levels have declined in most parts of the country while the reported incidence of asthma has risen[7], many researchers have begun to question the link between ambient air pollution and asthma (von Mutius, 2000a, 2000b; Vacek, 1999; Duhme et al., 1998). For example, the Committee on the Medical Effects of Air Pollution (Committee on the Medical Effects of Air Pollution, 2000) concluded that "overall evidence is small that non-biological outdoor air pollution has an important effect on the initiation and [provocation] of asthma", (2000). As a result, alternative theories have sprung up recently. One theory proposes that children are "too clean" because they often use antibiotics to combat minor illnesses. As a result, their immune systems do not develop properly and attack many harmless substances that enter the body (American Academy of Pediatrics, 2000). A second competing theory is that the changing lifestyles of children—poorer diets, less exercise, more time indoors - has led to the increase in asthma related illnesses (von Mutius, 2000a).

However, not all researchers have dismissed the role that pollution may play. There is a debate as to whether asthma rates have actually increased. Better detection of asthma

[4] Some representative studies include Desqueyroux and Momas (1999), Gouveia and Fletcher (2000), Fauroux (2000), Garty et al. (1998), Krupnick et al. (1990) and Norris et al. (1999).

[5] Criteria pollutants are non-toxic air pollutants considered most responsible for urban air pollution and are known to be hazardous to health. They include SO_2, NO_2, O_3, CO, PM_{10}, and lead.

[6] Other studies that have attempted to link pollution and general health use data that follow the same individuals over a short period of time to control for permanent health-related factors, such as smoking rates and exercise habits (Alberini and Krupnick, 1998; Portney and Mullahy, 1986, 1990). However, most of these studies focus on adults, and the results may not be directly applicable to children. Furthermore, a general limitation of these studies is that, given the limited number of observations over a short period of time, it is unlikely that there is enough variation in specific health outcomes, such as asthma, to obtain precise estimates.

[7] See footnote 2.

and different classifications of illness could explain some of the increases in individual and doctor reports. For example, what was long labeled wheezy bronchitis is now classified as asthma (Speizer, 2001). Recent expansions in Medicaid could also explain part of the increase in reported cases – as children's access to health care increases, there is a greater chance of early detection and treatment.

Many researchers have also questioned the methodological approaches used to identify the relationship between pollution and asthma (Nystad, 2000; Eggleston et al., 1999; von Mutius, 2000b; Bjorksten, 1999). Since air pollution is not randomly assigned, most studies have been largely unsuccessful in disentangling pollution from other confounding factors that affect health. Additionally, these studies do not account for direct responses to ambient levels of pollution. Furthermore, these studies tend to group all children into just one category, and we might expect a number of biological and behavioral factors to vary for children of different ages. Lastly, most studies conduct single pollutant analyses, which does not provide clear policy implications if pollutants are highly correlated.

A final reason to believe a connection between pollution and asthma might exist is that studies with more convincing empirical designs have found consistent effects of pollution on children's health. Chay and Greenstone (2003) use declines in pollution that resulted from the 1980–82 recessions and find a strong link between total suspended particles and infant mortality. Since most infant mortality is due to respiratory failure, it is reasonable to suspect that pollution could be related to other respiratory illnesses, such as asthma. Ransom and Pope (1995) use changes in pollution that resulted from the opening and closing of a steel mill due to a labor strike and find a large effect on bronchitis and asthma in children. Their study, however, does not identify the effect of specific pollutants, only the effect of the mill being opened or closed.[8]

3. Conceptual framework

One approach to understanding the impact of pollution on health would be to assume that everyone is unaware of the amount of pollution in the air. Therefore, ambient levels of pollution would serve as an unbiased proxy for an individual's exposure to pollution and pollution levels would not be correlated with any types of behavior. One could then estimate a relationship between health and pollution by regressing health outcomes on ambient levels of pollution as well as other exogenous factors that are related to both pollution and health, such as weather conditions.

However, this approach is oversimplified because individuals can undertake avoidance activities to reduce the effect of externalities, which makes an individual's exposure to pollution an endogenously determined variable.[9] This introduces two issues. First, there are many tools available to inform people when air pollution levels pose a threat to health.

[8] Another study (Friedman et al., 2001) that attempts to use a "natural experiment" caused by changing traffic patterns in Atlanta during the 1996 Olympics also does not identify the effects of particular pollutants. Moreover, this study does not consider the changing behavior of families in response to the Olympics in general.

[9] For a detailed description of avoidance (or averting) behavior, see Zeckhauser and Fisher (1976) or Bresnahan et al. (1997).

Home devices, such as peak expiratory flow (PEF) meters, can be used to measure lung functioning on a given day (if the individual already has a respiratory illness). California State law requires the announcement of air quality episodes, or "smog alerts", when pollution levels are predicted to exceed certain limits (Air Resources Board, 1990). State and local agencies are required to report a daily measure of air quality in large metropolitan areas, with newspapers a common source (U.S. EPA, 1999a). Many regional air quality offices, such as the California Air Resources Board, provide web pages with up-to-the-minute pollution details and e-mail notifications of dangerous pollution levels.[10] Many pollutants are directly visible – on high-smog days in Los Angeles, whitish clouds often cover the sky or a reddish-brown haze is visible around the horizon. If people directly respond to this information, then ambient pollution levels will not accurately represent their exposure to pollution.

A second issue arises because air quality, like many local public goods, is capitalized into housing prices, making it an attribute of a home that people can demand (Chay and Greenstone, 2000). Therefore, families with a higher value for cleaner air can locate in areas with better air quality.[11] These families may also make additional investments in their children's health – they may be less likely to smoke or more likely to seek preventative health care and have an existing relationship with a doctor. As a result, there are many confounding behavioral factors related to both pollution and health, making it difficult to identify the effect of pollution on health.[12]

Additionally, parents' investment decisions and pollution exposure may vary depending on the health stock or age of their child. This could occur because some children face a greater risk from the same exposure as other children or the costs to monitor behavior vary by child or age. For example, it is not uncommon for parents to insist on keeping tobacco smoke away from their infant, only to become more yielding as the child grows older, suggesting avoidance behavior might be more actively undertaken for younger children. Therefore, a specific child's exposure to pollution is potentially related to both the family specific endowment and type of care utilized.

These issues suggest that the ambient levels of pollution where a child lives are potentially related to many broadly defined factors that are difficult to fully observe, and omitting them from a regression could cause biased estimates of ambient pollution on asthma. I propose to control for these variables using the following innovations. First, I look at the effect of air pollution separately for children of different age groups. These groups correspond with both biological development and the type of care that families typically display towards children. I define the age categories of interest as follows: children age 0–1 (lung "branching" occurring at rapid rate; infants most protected by parents and most likely to use hospital for illness); 1–3 (alveoli develop and mature; children spend more time in day care); 3–6 (children more likely to enroll in preschool/kindergarten); 6–12 (elementary school); and

[10] For example, visit http://www.epa.gov/airnow/ to find daily pollution levels throughout the United States.

[11] Families do not need to have direct preferences for this attribute. However, because air quality is an input in the health production function, people with preferences regarding health will have implicit tastes for air quality.

[12] This is analogous to the confounding that arises in estimating the effect of school quality on test scores. Parents who choose to live in areas with better school quality may also make additional investments in their children, making it difficult to identify the effect of school quality.

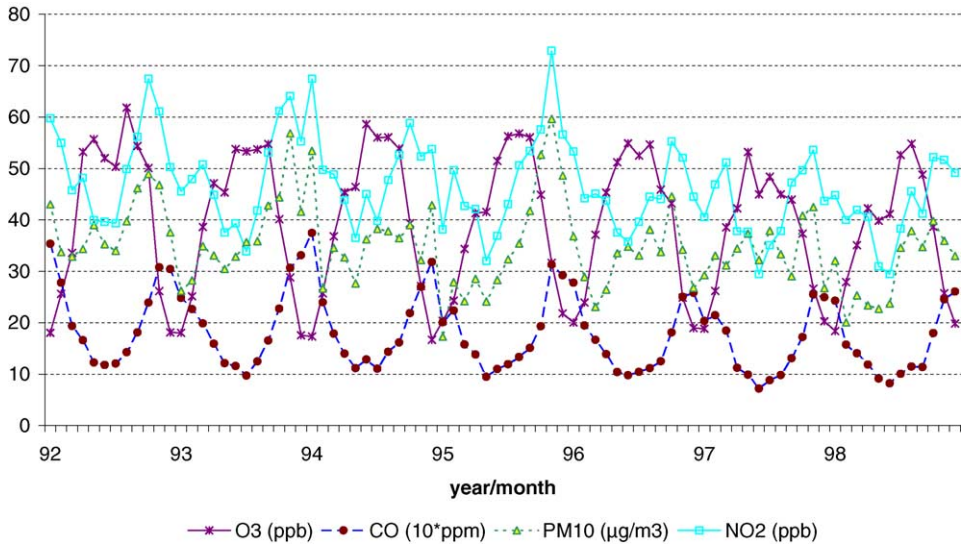


Fig. 1. Seasonal variation in pollution.

12–18 (secondary school). This will allow for different potential biological and behavioral responses to pollution by the age of the child.

Second, by creating monthly time-series data at the zip code level, I define the unit of observation as the zip code/month and specify a zip code-year fixed effect (FE). This will capture permanent observed and unobserved factors within a zip code-year that affect health, such as average smoking rates, average indoor pollution levels, and average health care decisions to the extent that they are constant within a year or do not change in ways that are correlated with pollution. Since the zip code is a finely defined geographic area with frequent social interactions amongst residents, the FE will capture a large share of potentially omitted characteristics.

The third innovation comes from using the diverse seasonal variation in pollution in California that arises from local microclimates and geography. While it is plausible that there are seasonal changes in health behavior that are correlated with changes in pollution, the key factor is that these seasonal variations in pollution are different throughout California depending on the unique physical characteristics of each area. For example, levels of ozone increase in the summer at a greater rate because ozone is formed in the presence of sunlight. Particulate matter is trapped by fog in winter weather. CO levels increase in cold, stagnant weather. Fig. 1, based on zip code level pollution estimates (described in more details below), shows the strong seasonal patterns of these pollutants. Furthermore, ozone increases at a greater rate in the summer in hotter and sunnier areas, such as southern and central California. PM_{10} increases in drier areas in the summer and fall, but increase in colder areas in the winter because of increased use of combustion sources (Nystrom, 8/2001). To highlight some of this diversity, Fig. 2 shows monthly pollution levels for coastal counties in southern California, an area where we might expect similar seasonal variations in health

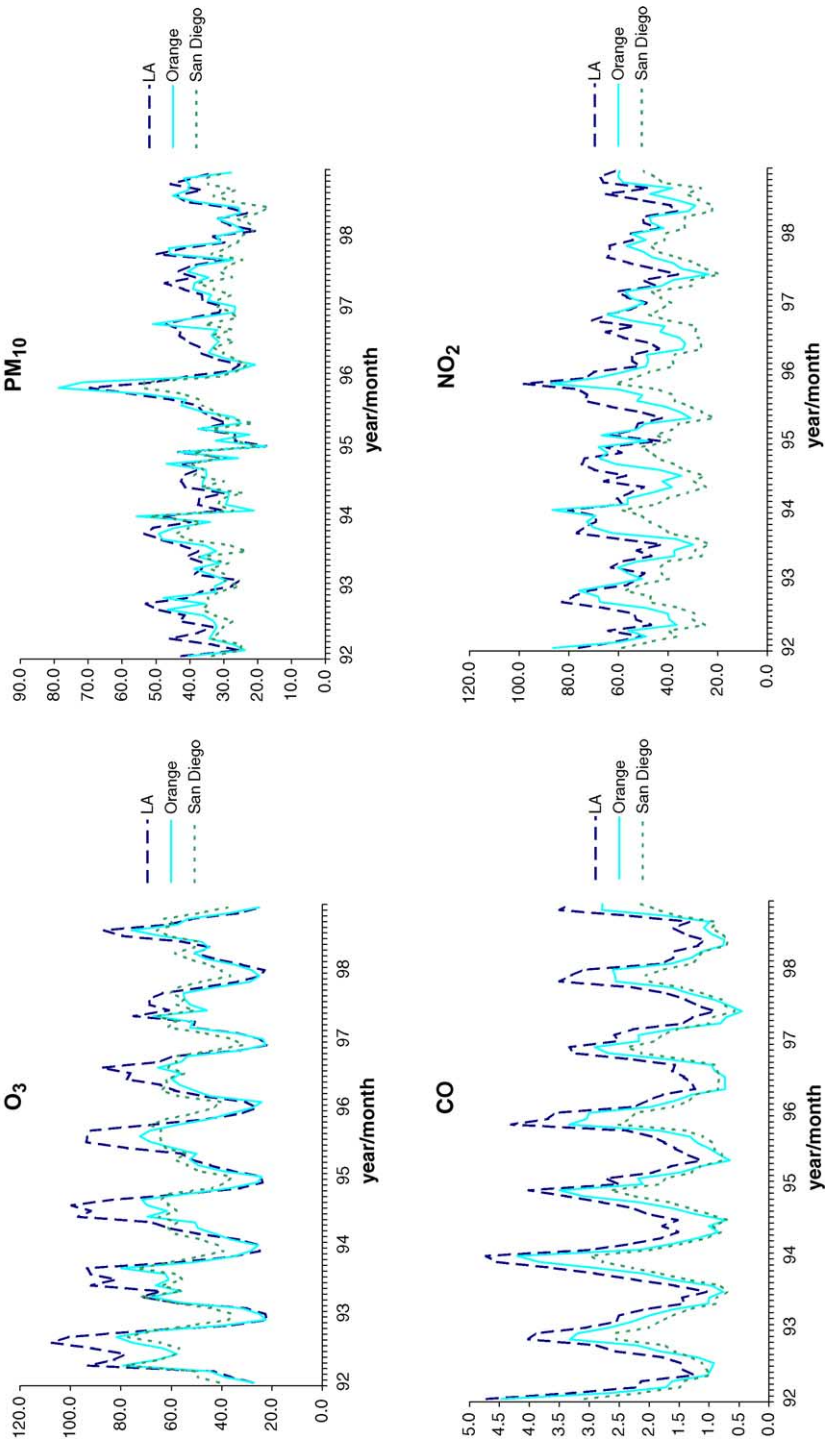


Fig. 2. Seasonal variation in pollution by county.

behavior and face comparable weather patterns. Focusing on ozone, Los Angeles and Orange counties have comparable levels in the winter, but Los Angeles experiences higher levels in the summer. Meanwhile, San Diego experiences higher winter levels that Los Angeles and Orange, but lower summer levels. Since these patterns in pollution vary throughout California and are naturally occurring, it is reasonable to assume that they are independent of many seasonal investments in health.

In sum, I will compare how seasonal changes in pollution within a given zip code-year affect changes in seasonal asthma rates for a specific age group.[13] The following example of smoking rates and outdoor pollution highlights how the empirical strategy works. Failing to control for smoking is only a problem if smoking behavior is related to both pollution and asthma. By looking at separate age groups, I circumvent the need to control for how parents monitor tobacco smoke around their children based on the age of the child. By using zip code-year fixed effects, I look at whether changes in pollution are linked to changes in asthma within a zip code for each year. If smoking either does not change with changes in pollution, or if it changes in a way that is unrelated to changes in pollution, then the fixed effect would control for smoking behavior. Smoking behavior, however, may change within a year - people may be more likely to smoke when they spend more time outside as the weather improves. If this is the case, the fixed effect will not capture the changing smoking patterns. However, if smoking patterns do not change from one season to the next in a way that is correlated with the seasonal changes in pollution unique to that area, then I will not need to explicitly control for smoking behavior.

While this identification strategy overcomes many problems, there is one main source of endogeneity that remains—contemporaneous avoidance behavior. Since people can directly respond to daily pollution, this will not be captured by the identification strategy. Although I include some measures of avoidance behavior, these measures only capture part of avoidance behavior and only as it relates to ozone. However, if contemporaneous avoidance behavior is positively related to pollution levels and avoidance behavior lowers the likelihood of having an asthma attack, omitting it will yield a lower bound of the true effect.

To proceed with estimation, I estimate the following model separately for each age group:

$$_{zyt} = \ _{0}\ _{zyt} + \ _{1}\ _{zyt} + \ _{2}\ _{zyt} + \ _{zy} + \ _{yt} + \ _{zyt} \tag{1}$$

where the subscripts z, y, and t indicate zip code, year, and month, respectively, *Y* is the number of ER asthma admissions, *P* are ambient air pollution levels, *A* is contemporaneous avoidance behavior that directly affects the child's exposure to pollution, *X* are other factors that affect health (such as weather), and is an i.i.d. error term. The key identification parameters are $_{zy}$, zip code-year fixed effects, and $_{yt}$, year-month dummy variables. $_{0}$ is the coefficient vector of interest. The main hypothesis to test is whether $_{0} = 0$, namely that pollution has no effect on asthma admissions.

[13] One notable limitation of using seasonal changes in pollution is that, by smoothing out daily variation, some valuable information may be lost. However, this only affects the efficiency of the estimates unless there are threshold effects such that pollution beyond a certain level induces asthma attacks. Additionally, using seasonal variation will not provide evidence on long-term health effects.

4. Data

4.1. Sources

The California Hospital Discharge Data (CHDD) is a rich source of individual health outcomes. This data set records the principal diagnosis of the patient upon release from the hospital[14], the month of admission,[15] the zip code of residence, as well as the sex, race, age, and the expected source of payment for all individuals discharged from a hospital in the state of California. Data are available from 1992 to 1998 and each year contains on average over 800,000 hospital discharges for children under age 18 (not including newborns).

While hospital data does not include information on all asthma attacks, the CHDD offers three key advantages over self-reported surveys. First, hospital discharges, in particular ER admissions, are a more objective measure of asthma and are less likely to be sensitive to reporting biases.[16] Second, there are a large number of observations available each year in the CHDD. Third, having the zip code of the patient enables me to specify a zip code fixed effect and to merge other key data sources at the zip code level.

The key data merged are atmospheric pollution levels from Environmental Protection Agency (EPA) air monitoring stations throughout California. The monitor data are readily available from 1982 until the present and are the most detailed data recording ambient levels of criteria pollutants. Furthermore, they contain the exact location of the monitor, enabling them to be merged with the CHDD. Fig. 3 shows O_3 monitors in California in 1999 along with county outlines. These monitors are mainly located in the more densely populated areas (shaded in gray). Fig. 4 highlights Los Angeles County, showing again O_3 monitors and now the outlines of zip codes. Since Los Angeles is a diverse county both demographically and geographically and there are many monitors to capture local pollution levels, assigning pollution at the zip code level should produce more reliable measures than from assigning it at a broader level.

I also merge other data sources at the zip code level. Monthly meteorological data from the National Climatic Data Center contains various measures from more than 1000 weather stations in California as well as their exact location.[17] The California Association of Realtors provides monthly zip code level information on the number of homes and average and median sales price from 1991 to the present. Using 1990 Census estimates of population counts by age for each zip code and annual county estimates by age from the Demographic Research Unit of the California Department of Finance, I have approximated the annual population for each zip code and age group.

As proxies for avoidance behavior, I merge the number of smog alerts announced in each month. Air quality episodes, or "smog alerts", are required by California law to be issued by

[14] This is assigned according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD- 9-CM) by the U.S. Department of Health and Human Services.

[15] The exact day of the month is censored in the version of the data that has already been released to me. Only an indicator for the day of the week is available.

[16] ER admissions account for approximately 67% of all hospital admissions for asthma. When performing the analysis using all hospital admissions, the results did not change considerably.

[17] The meterological data are merged using the same inverse-distance weighted technique used to approximate zip code levels of pollution (described below).

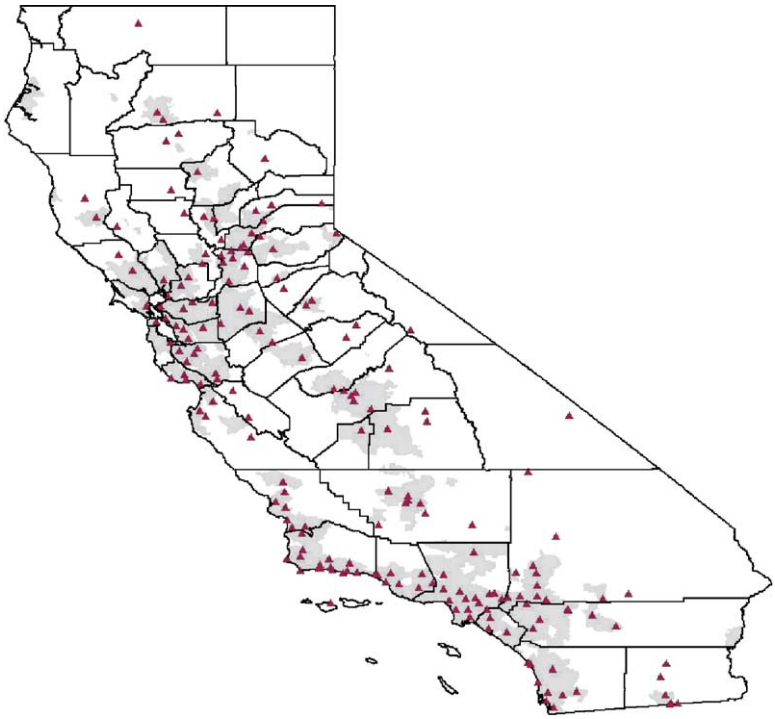


Fig. 3. Ozone monitors in California.

local air quality management districts (AQMD)[18] when criteria pollutants exceed levels as specified by the California Air Resources Board.[19] When this occurs, schools are directly contacted and are urged to limit physical activities for children until pollution levels ease, while other sensitive people are advised to avoid the pollution by remaining indoors (Air Resources Board, 1990). While these advisories are required to be announced for all of the criteria pollutants, historically announcements have only be made for ozone levels, and as a result the advisories are commonly referred to as "smog alerts."

4.2. Linking pollution

To approximate a monthly time-series of pollution at the zip code level, I first calculated the coordinates for the centroid of each zip code in California. Using the reported coordinates of the EPA monitors, I then measured the distance between each centroid and each monitor. Finally, I calculated the level of pollution for a zip code by averaging reported values from all monitors within 20 miles of the centroid, weighting by the inverse of the distance from

[18] There are currently 17 air quality management districts in California.

[19] While I only possess this data for the AQMD that covers Los Angeles, Orange, Riverside and San Bernardino counties, it is unlikely that any other area has experienced smog alerts during this period.

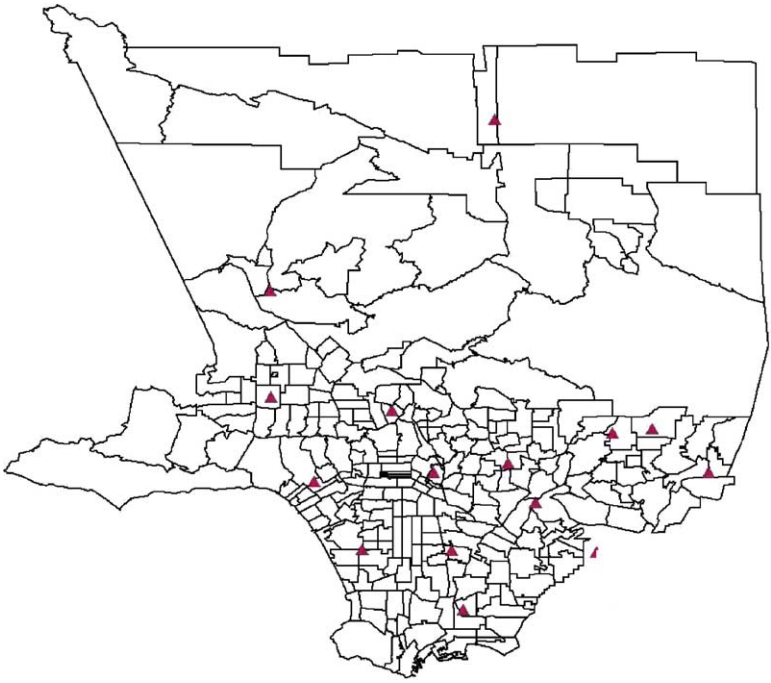


Fig. 4. Ozone monitors in Los Angeles county.

the centroid to the monitor. Therefore, I define pollution in zip code z at time *yt* as:

$$_{zyt} = \frac{(\ _{jyt} \times 1\ (\quad | \quad \leq 20))}{1\ (\quad | \quad \leq 20)} \tag{2}$$

where D_j is the distance from monitor *j* to the centroid of zip code z and P_{jzt} is the pollution measure at monitor *j* in year y in month t.

Five immediate issues arise in measuring pollution in this way. First, the choice of 20 miles as the cutoff is arbitrary. To test the sensitivity of this assumption, I also assigned pollution levels using distance cutoffs of 10 and 5 miles. Panel A of appendix Table A.1 shows the correlation between pollution levels calculated using the various cut-offs. These correlations are remarkably high, with none below 0.95, suggesting that the choice of distance is unlikely to introduce any biases.[20]

Second, many monitors have been added or removed over the time period studied. This occurs because pollution monitors are installed in areas where pollution exceeds NAAQS, and can also be removed from an area if it falls below NAAQS (U.S. EPA, 1999b). As a result, monitors are more likely to be placed in areas where pollution levels have been increasing, and less likely to exist in areas where pollution has been declining. To assess the implication

[20] When performing the analysis using a distance of 10 or 5 miles, the results were nearly identical, consistent with these high correlations.

of this, I estimate (2) in two ways: using all monitors from 1992 to 1998 and using only continuously operated monitors from 1992 to 1998. Panel B of appendix Table A.1 shows the number of monitors over time for both methods and the correlation between monthly zip code levels of each pollutant calculated by each method. The overall number of monitors has not changed considerably and the correlations for all are at least 0.98, indicating that the sampling technique used for monitors should not interfere with inference.

Third, there are many factors that affect how pollutants travel, such as wind, rain, and the size of the pollutant particle, and this may affect how well (2) measures the actual pollution concentration[21]. For example, particulate matter, such as PM_{10}, settles to the ground at a much quicker rate than do gaseous pollutants (Wilson and Spengler, 1996). To get a sense of how accurate the above approach is, I estimate the level of pollution at each monitor (as opposed to zip code) using the above formula as if the monitor of interest were not there. Therefore, I estimate the amount of pollution at a given monitor based on the pollution levels at monitors less than 20 miles away. I do this for all monitors and then calculate the correlation between the estimated pollution and the actual pollution, shown in panel C of appendix Table A.1. The correlations for O_3 and NO_2 are remarkably high. This is not surprising for O_3 given it is formed in the atmosphere, as opposed to being directly emitted. For PM_{10} and CO, the correlations are slightly lower, but still quite high at over 0.75. This indicates the reasonable accuracy of this method for assigning pollution and suggests it does not appear to be a major concern for this analysis.

Fourth, while it is crucial to control for multiple pollutants simultaneously, trying to separately identify the effect of each pollutant can be difficult if pollutants are highly correlated. Many pollutants originate from similar sources, as the preceding chart indicated. Appendix Table A.2 shows the correlation matrix for the pollutants as assigned according to (2). O_3 does not appear highly correlated with any other pollutants, while NO_2 appears highly correlated with CO and PM_{10}. This may make it difficult to obtain precise estimates for NO_2, making it useful to analyze models with the pollutants included individually.

Fifth, since monitors tend to exist in more polluted and populated areas, it is important to understand how the characteristics of the population in these areas differ from those that are excluded from the analysis. Appendix Table A.3 shows various demographic characteristics for zip codes that are within 20 miles of a monitor for each of the pollutants and zip codes that are not. While all of the variables shown are statistically different, the driving force behind these differences appears to be the percent of the population of the zip code that lives in urbanized areas. This coincides with the monitor locations shown in Fig. 3. Although this limits the representativeness of these findings, rural areas represent a much lower fraction of the population and are less likely to experience high levels of pollution.

4.3. Trends and descriptive statistics

Table 1 panel A shows the descriptive statistics of the data used in the analysis, including the "between" and "within" zip code-year variation of each variable.[22] For the pollutants,

[21] While I obtained measures of precipitation to include in the analysis, wind data is not as widely available. Furthermore, there is much debate on how to incorporate wind data.

[22] The "between" standard deviation is calculated using $^{-}$ and the "within" is calculated using $= {}^{-} + {}^{-}$.

Table 1
Summary statistics

A. Summary statistics

	Observations	Groups	Mean	S.D.	'Between' zip-year S.D.	'Within' zip-year S.D.
O_3 (ppb)	54,963	4991	38.905	17.803	9.830	14.855
CO (ppm)	54,963	4991	1.777	1.037	0.679	0.790
PM_{10} (g/m^3)	54,963	4991	34.210	13.984	9.874	9.933
NO_2 (ppb)	54,963	4991	45.947	17.171	14.392	9.539
ER asthma rate age 0–1	54,843	4979	0.431	1.572	0.964	1.382
Population age 0–1	54,843	4979	578	390	391	0
ER asthma rate age 1–3	43,624	4949	0.135	0.755	0.469	0.668
Population age 1–3	43,624	4949	2347	1449	1416	0
ER asthma rate age 3–6	44,462	4959	0.167	1.066	0.762	0.916
Population age 3–6	44,462	4959	2159	1327	1305	0
ER asthma rate age 6–12	50,602	4987	0.088	0.393	0.313	0.344
Population age 6–12	50,602	4987	3568	2209	2221	0
ER asthma rate age 12–18	55,377	4990	0.072	0.411	0.202	0.380
Population age 12–18	55,377	4990	1812	1159	1170	0
% Normal neonates	54,963	4991	0.698	0.078	0.075	0.020
% Government insurance	54,962	4990	0.399	0.216	0.216	0.023
Ave. max. temperature (°F)	54,963	4991	0.073	0.011	0.004	0.010
Total precipitation (in.)	54,963	4991	0.007	0.011	0.003	0.010
Semi-annual house price	45,662	4194	20,007	12,197	11,841	5.503
Smog alerts	54,963	4991	0.288	1.582	0.808	1.353

B. Pollution and asthma by SES

Pollutant	High	Low	Age	High	Low
O_3	38 283 (0.123)	40 117 (0.138)	0–1	0 493 (0.011)	0 729 (0.015)
CO	1 652 (0.007)	1 953 (0.008)	1–3	0 177 (0.004)	0 179 (0.009)
PM_{10}	31 851 (0.095)	38 160 (0.104)	3–6	0 193 (0.004)	0 235 (0.012)
NO_2	42 964 (0.118)	50 361 (0.123)	6–12	0 102 (0.002)	0 127 (0.004)
Observations	19,458	19,299	12–18	0 122 (0.003)	0 130 (0.005)

Note: The "between" standard deviation is calculated using $^{-}$ and the "within" is calculated using $^{-} - ^{-} + ^{-}$. Standard errors in parenthesis. Low SES is defined as zip code percentage of high school dropouts greater than the median level of high school dropouts.

it is not unusual for the within zip code variation to exceed the between zip code variation, as is the case for O_3 and CO. For asthma admission rates,[23] younger children have a greater likelihood of visiting the ER,[24] with infants approximately six times more likely to visit the ER than children over 6 and 1–6 year-olds two times more likely to visit than children over 6. Most of the variation in asthma rates comes from within the zip code-year. The patterns in variation for asthma and pollution suggest ample variation for obtaining precise estimates using the identification strategy described above.

Table 1 panel A also shows variables that represent A_{zyt} and X_{zyt}. House prices are designed to reflect changes in asset wealth and are a "sufficient" statistics for many demographics of a given area, such as school quality and crime rates. The percentage of newborns with government-sponsored health insurance (calculated from the CHDD) is used as a measure of changes in the bottom of the income distribution.[25] The percentage of normal newborns (calculated from the CHDD[26]) is used to approximate the health stock for infants. Monthly hospital admissions for influenza are included to control for co-morbidities that may be inducing asthma episodes, rather than pollution itself. Average maximum temperature and inches of precipitation both affect the likelihood of being outdoors and may directly exacerbate asthma symptoms (American Lung Association, 8/2001).

Since asthma disproportionately attacks children of low SES, Table 1 panel B shows pollution levels and ER asthma rates for two SES groups. I define SES groups as above and below the median for the percent of adults over 25 years old in a zip code without a high school diploma. The average levels of all pollutants are higher for the low SES groups. Asthma rates for low SES are almost twice as high as high SES for infants, but are only slightly higher for children over age 1. These differences in pollution and asthma rates by SES are statistically significant, except for asthma for age 1–3.[27]

In turning to annual trends, Fig. 5 shows asthma patterns in California and other regions in the United States using the National Hospital Discharge Survey (NHDS).[28] The northeast has the highest admission rate, followed by the Midwest, the south, and then the west. The pattern for California is similar to that for the entire United States, but at a level that is almost 50% lower. In turning to monthly patterns, Fig. 6 shows asthma patterns over time for each age group separately. Immediately evident are the strong seasonal patterns for admissions for all age groups. Rates for each age group increase on average anywhere from two to three times from the lowest month to the highest. Furthermore, the seasonal patterns differ across

[23] Asthma is labeled as ICD-9-CM 493.

[24] ER admissions are distinguished from other admissions according to the "source of admission" variable from the CHDD.

[25] There was only one expansion in medicaid eligibility that affected newborns during the time period studied. In February of 1995, eligibility was extended from 185 to 200% of the federal poverty level. Although Access to Infants and Mothers (AIM) also increased during this period, less than 0.6% of all births in California are paid for by AIM (Managed Risk Medical Insurance Board, 8/2001).

[26] In the CHDD, newborns are classified into one of the following seven categories: (1) died or transferred (2) extreme immaturity or respiratory distress syndrome (3) prematurity with major problems (4) prematurity without major problems (5) full term with major problems (6) neonate with other significant problems and (7) normal newborn.

[27] These patterns are also present when SES is defined by race or income.

[28] The NHDS does not provide information to separately identify emergency and non-emergency hospital admissions and the only geographic identifier is the region.

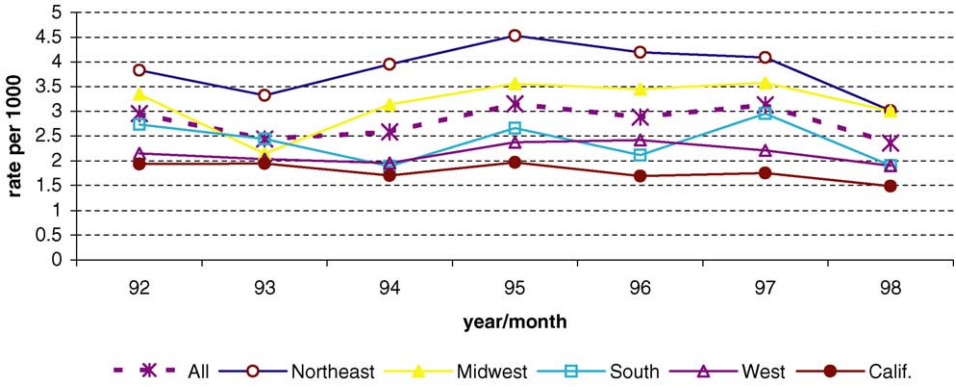


Fig. 5. All hospital admissions for asthma for children in U.S.

age groups. The high season for infants is the winter, whereas high season for teens is the fall. These striking patterns demonstrate the importance of looking at age groups separately and the potential value in exploiting seasonal variation.

Before turning to the estimation, a case study of a specific zip code highlights the main findings of this analysis. Fig. 7 plots monthly standardized pollution levels and asthma counts for children ages 3–6 in zip code 92335 (Fontana in San Bernardino County)). A strong pattern between asthma and CO emerges, with peaks and trough occurring at roughly the same time throughout the entire time period. Asthma and O_3 appear negatively correlated, with O_3 peaking in the summer. While at times asthma follows the patterns of PM_{10} and NO_2, the pattern tends not to persist for the entire time period, indicating a potential link between CO and asthma.

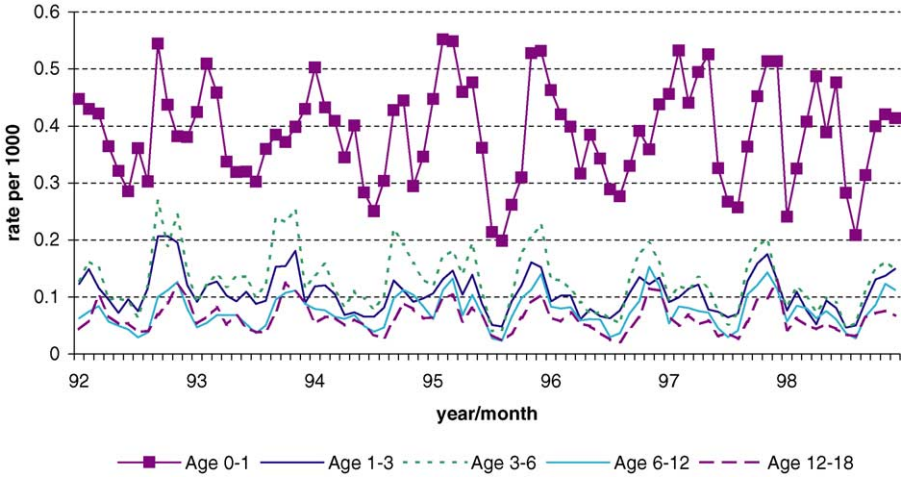


Fig. 6. Monthly ER asthma rates by age.

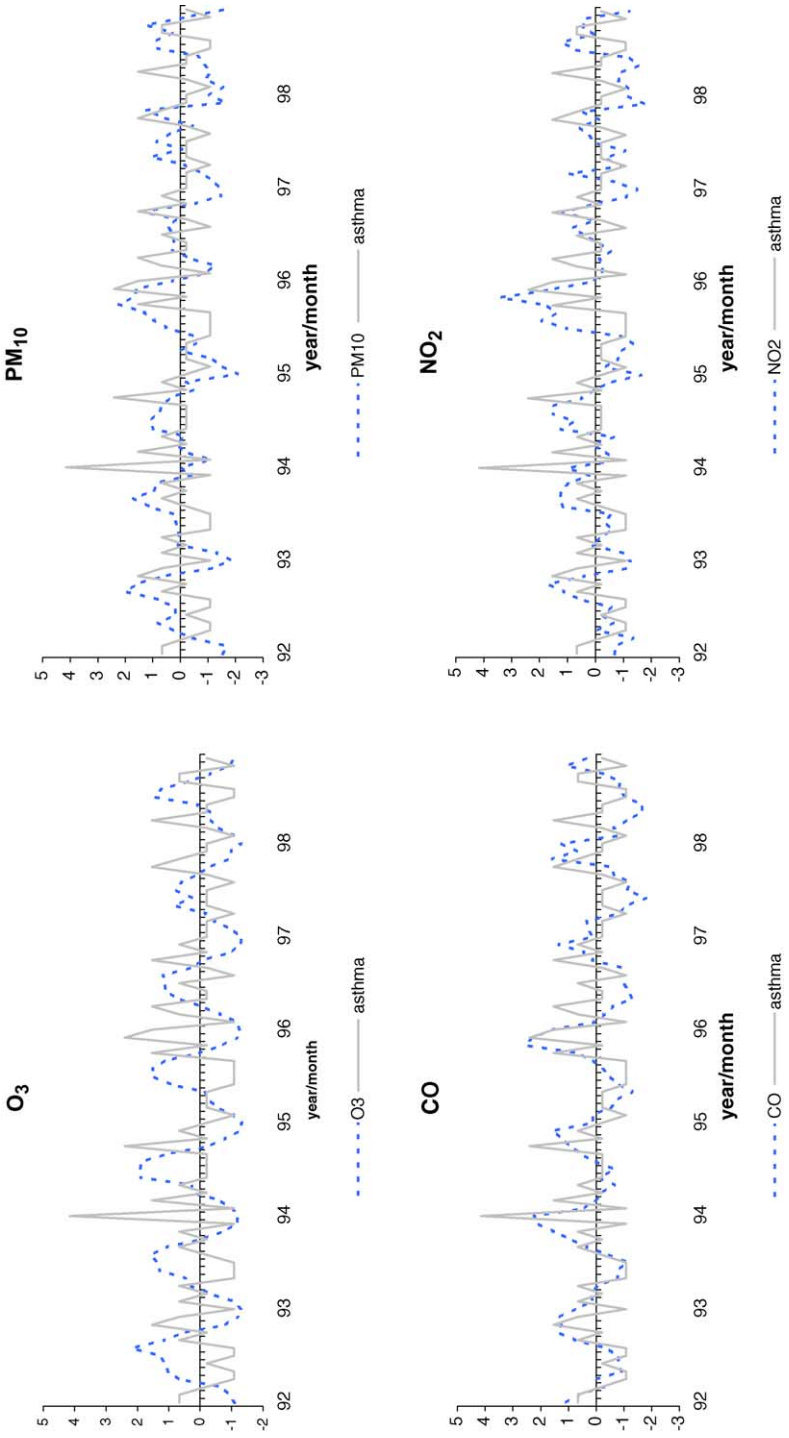


Fig. 7. Polluton and asthma for ages 3-6 in Zip 92335.

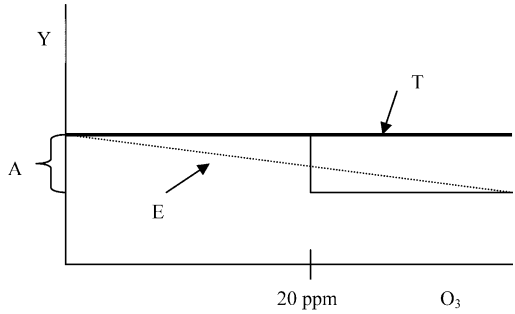
5. Results

5.1. Main results

The first set of results, fixed effect estimates of Eq. (1) without any direct controls for avoidance behavior, indicate that pollution has a differential impact on infants as compared to older children. As indicated in Table 2 panel A,[29] none of the pollutants are significantly related to asthma ER hospitalizations for infants. However, for all older age groups, CO is positive and significantly correlated with asthma. One explanation for the difference across age groups is that asthma is often difficult to precisely identify in infants because of communication limitations, little history of respiratory illnesses, and birth complications (Letourneau et al., 1992). O_3 has a negative effect on admissions, and is precisely estimated for ages 6–12. NO_2 and PM_{10} are generally positively correlated with asthma, but only NO_2 is significant for Ages 6–12.

In terms of the control variables, temperature and precipitation are negatively correlated with asthma. While weather may have a direct effect on asthma, such as precipitation "cleaning" the air (Wilson and Spengler, 1996), weather may indirectly affect asthma by altering children's exposure to pollution by changing the amount of time spent inside. Influenza has a positive although insignificant effect on asthma admissions, which is consistent with co-morbidity theories. The coefficients for the demographic variables are almost always imprecisely estimated. Since these variables often have significant effects on health outcomes, this suggests that the fixed effects appear to control for a large amount of observed as well as unobserved heterogeneity.

While the presence of negative coefficients for O_3 in Table 2 panel A may at first seem surprising, not controlling for contemporaneous avoidance behavior under-estimates the true effect if avoidance behavior increases as pollution increases. Furthermore, if people respond to an increase in pollution by increasing avoidance behavior to the point that health actually improves, it can induce a negative effect. The following diagram illustrates how negative effects could arise for O_3.



T = true dose-response
E = estimated dose-response
A = avoidance behavior

[29] For ease of interpretation, all pollutants have been standardized to have a mean of zero and standard deviation of one.

Table 2
Main results

Panel A. Fixed effect estimates by age group

	(1) Age 0–1	(2) Age 1–3	(3) Age 3–6	(4) Age 6–12	(5) Age 12–18
O_3	−0 007 (0.010)	−0 018 (0.011)	−0 016 (0.011)	−0 054** (0.010)	−0 010 (0.008)
CO	−0 010 (0.010)	0 024* (0.011)	0 049** (0.011)	0 023* (0.011)	0 021* (0.009)
PM_{10}	−0 001 (0.008)	−0 004 (0.009)	0 000 (0.009)	0 016 (0.008)	0 003 (0.007)
NO_2	0 009 (0.014)	0 002 (0.016)	0 006 (0.016)	0 041** (0.015)	0 005 (0.013)
Ave. max. temp./10,000	−3 690** (1.030)	−4 134** (1.163)	−1 885 (1.133)	−2 260* (1.053)	−0 512 (0.900)
Total precip./10,000	−0 522 (0.777)	−2 257** (0.876)	−2 717** (0.916)	−2 938** (0.827)	−1 133 (0.706)
% Normal neonates	0 093 (0.154)	n/a	n/a	n/a	n/a
Log (house price/10,000)	−0 016 (0.026)	0 025 (0.031)	−0 024 (0.030)	−0 006 (0.030)	−0 018 (0.027)
% Government insurance	−0 057 (0.119)	−0 377* (0.149)	0 104 (0.146)	−0 166 (0.137)	0 052 (0.115)
Influenza admissions	0 014 (0.018)	0 008 (0.036)	0 046 (0.047)	0 068 (0.048)	0 028 (0.040)
Observations	38,779	33,037	34,528	38,778	31,758
Number of groups	3372	3360	3490	3602	2759
R-squared	0 28	0 24	0 25	0 28	0 15

Panel B. Fixed effect estimates by age group with controls for avoidance behavior

	(1) Age 0–1	(2) Age 1–3	(3) Age 3–6	(4) Age 6–12	(5) Age 12–18
O_3	−0 005 (0.011)	−0 006 (0.012)	−0 006 (0.012)	−0 045** (0.011)	−0 009 (0.009)
CO	−0 010 (0.010)	0 027* (0.011)	0 051** (0.011)	0 025* (0.011)	0 021* (0.009)
PM_{10}	−0 001 (0.008)	−0 005 (0.009)	−0 001 (0.009)	0 015 (0.008)	0 003 (0.007)
NO_2	0 009 (0.014)	0 002 (0.016)	0 006 (0.016)	0 042** (0.015)	0 005 (0.013)
# of smog alerts	−0 001 (0.002)	−0 007** (0.003)	−0 005* (0.003)	−0 005* (0.002)	−0 000 (0.002)
Observations	38,779	33,037	34,528	38,778	31,758
Number of groups	3372	3360	3490	3602	2759
R-squared	0 28	0 24	0 25	0 28	0 15

Robust standard errors in parenthesis. Pollutants are normalized to have a mean of zero and standard deviation of one. All columns include year/month dummy variables and an indicator if house price information is missing. The sepcification in each column is identical to the corresponding column panel A of Table 2.

* Significant at 5%.

** Significant at 1%.

When ozone exceeds 20 ppm, a smog alert is announced. If schools or parents respond by keeping their children inside, children may exercise less as a result. Since exercise is believed to induce asthma and is not directly observed, by omitting *A* I would estimate line *E* instead of *T*, yielding a spurious negative effect of O_3 on asthma hospitalizations. Although this diagram assumes that ozone has no effect on asthma, the same negative effect could occur if ozone has a positive effect on asthma.

To test the impact from omitting avoidance behavior, I add to the model the number of smog alerts announced in each month.[30] Since smog alerts are only announced with respect to O_3, this only tests how estimates for O_3 changes. The results from including this variable, reported in Table 2 panel B, show that smog alerts have a negative effect on asthma admissions for ages 1–12, supporting the notion that avoidance behavior is actively undertaken. Meanwhile, the negative effect for O_3 becomes considerably smaller and there are no qualitative changes in the other pollutants. If O_3 has no effect on asthma admissions, in order for smog alerts to have a negative effect on asthma admissions children must be doing something in addition to avoiding pollution that reduces the likelihood of having an asthma attack, such as exercising less.[31] However, it is possible that additional controls for avoidance behavior with respect to O_3 could uncover a positive effect of O_3 on asthma hospitalizations.[32] These results indicate that omitting avoidance behavior induces a lower bound of the biological effect of pollution on health.

An additional concern with these estimates is that by controlling for pollutants simultaneously it may be difficult to separately identify the effects of each pollutant. For example, as mentioned in the data section, NO_2, CO, and PM_{10} are highly correlated pollutants. Table 3 shows estimates from single-pollutant models. The effects of CO and NO_2 are both larger in general. If automobile exhaust is truly a contributing factor to asthma, this finding is not surprising because automobile exhaust contributes heavily to both pollutants.

An important issue to explore is whether the effect of pollution on asthma is the same for all subsets of the population. Some groups, such as children with low health stock, may face different risks from comparable levels of pollution. Additionally, as Table 1 panel B shows, pollution levels and asthma rates are higher for children of low SES. To assess the importance of this, I interact each variable in Eq. (1) with an SES indicator as defined in Table 1 panel B.[33] These results, shown in Table 4, indicate that both O_3 and CO have a larger effect on low SES children, with positive and significant effects on the interaction term for ages 3–6 and 12–18. This supports the "double jeopardy" hypothesis that low SES children are not only exposed to higher levels of pollution but also are more harmed by similar amounts of pollution.

[30] Since it is possible that the marginal effect of smog alerts varies with the number of smog alerts announced, I also estimated a model that included the square of smog alerts. This term was insignificant and did not affect the other estimates.

[31] An alternative interpretation of these results is that smog alerts proxy for high levels of O_3. This interpretation suggests that O_3 levels below 20 ppm have no effect on asthma admissions but levels above 20 ppm reduce the number of admissions, which is an unlikely scenario.

[32] Additionally, including avoidance behavior controls specific to the other pollutants can increase the magnitude of the coefficients for those pollutants.

[33] I also performed this analysis by defining SES according to race and income, and the results were comparable.

Table 3
Fixed effect estimates with single-pollutants

Pollutants	(1) Age 0–1	(2) Age 1–3	(3) Age 3–6	(4) Age 6–12	(5) Age 12–18
O_3	−0 001 (0.009)	−0 014 (0.010)	−0 015 (0.010)	−0 029** (0.009)	−0 01 (0.008)
CO	−0 007 (0.009)	0 027** (0.009)	0 053** (0.010)	0 047** (0.009)	0 025** (0.008)
PM_{10}	−0 001 (0.007)	−0 002 (0.007)	0 007 (0.008)	0 019** (0.007)	0 006 (0.006)
NO_2	0 001 (0.011)	0 009 (0.013)	0 028* (0.013)	0 047** (0.012)	0 015 (0.010)
Observations	38,779	33,037	34,528	38,778	31,758
Number of groups	3372	3360	3490	3602	2759

See notes to Table 2B. Each row within an age group represents results from regression with only that pollutant.

* Significant at 5%.

** Significant at 1%.

Table 4
Fixed effect estimates by age group and SES

	(1) Age 0–1	(2) Age 1–3	(3) Age 3–6	(4) Age 6–12	(5) Age 12–18
O_3	−0 007 (0.013)	−0 003 (0.015)	−0 038** (0.014)	−0 044** (0.013)	−0 022 (0.011)
CO	−0 017 (0.013)	0 000 (0.014)	−0 016 (0.014)	0 000 (0.014)	−0 003 (0.012)
PM_{10}	−0 006 (0.010)	−0 012 (0.011)	0 004 (0.011)	0 011 (0.010)	0 005 (0.009)
NO_2	0 021 (0.017)	−0 001 (0.020)	0 020 (0.020)	0 040*	