NBER WORKING PAPER SERIES

DOES POLLUTION INCREASE SCHOOL ABSENCES?

Janet Currie
Eric Hanushek
E. Megan Kahn
Matthew Neidell
Steven Rivkin

Working Paper 13252
http://www.nber.org/papers/w13252

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2007

We thank Alfonso Flores-Lagunes, Matthew Kahn, Walter Nicholson, and Jessica Reyes for helpful comments as well as participants at the 2005 AEA meetings and the 2006 SOLE meetings. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Does Pollution Increase School Absences?
Janet Currie, Eric Hanushek, E. Megan Kahn, Matthew Neidell, and Steven Rivkin
NBER Working Paper No. 13252
July 2007
JEL No. I18,Q51

ABSTRACT

We examine the effect of air pollution on school absences using unique administrative data for elementary and middle school children in the 39 largest school districts in Texas. These data are merged with information from monitors maintained by the Environmental Protection Agency. To control for potentially confounding factors, we adopt a difference-in-difference-in differences strategy, and control for persistent characteristics of schools, years, and attendance periods in order to focus on variations in pollution within school-year-attendance period cells. We find that high levels of carbon monoxide (CO) significantly increase absences, even when they are below federal air quality standards.

Janet Currie
International Affairs Building
Department of Economics
Columbia University - Mail code 3308
420 W 118th Street
New York, NY, 10027
and NBER
jc2663@columbia.edu

Eric Hanushek
Hoover Institution
Stanford University
Stanford, CA 94305-6010
and NBER
hanushek@stanford.edu

E. Megan Kahn
Dept. of Economics
Amherst, MA 01002-5000
emkahn@amherst.edu

Matthew Neidell
Department of Health Policy & Management
Mailman School of Public Health
Columbia University
600 W 168th St., 6th floor
New York, NY 10027
mn2191@columbia.edu

Steven Rivkin
Amerst College
Department of Economics
P.O. Box 5000
Amherst, MA 01002-5000
and NBER
sgrivkin@amherst.edu

I. Introduction

Even though substantial policy initiatives are aimed at reducing air pollution, uncertainty remains about the nature and extent of benefits from these actions. Existing epidemiological studies point to a variety of health impacts, but it remains difficult to assess the economic or social valuations for these. It is also difficult to be confident that these studies separate the causal impacts of pollution from correlated effects of neighborhoods, poverty, and a variety of household choices. We focus on how pollution affects school absences. By matching detailed schooling records with variations in the level of specific pollutants, we are able to establish a strong link to school absences.

A large literature links child health and human capital attainment. Grossman and Kaestner (1997) summarize this literature and point to school absences as a major causal link in this relationship—children who miss a lot of school achieve poorer grades, are less engaged with school, and are more likely eventually to drop out. Absences are also of concern to parents who have to miss work and to school districts because state funding frequently depends on attendance. In some states, schools can lose as much as $50 per day per unexcused absence. Nonetheless, policy interventions that might reduce absences are less clear.

Air pollution is a possible cause of school absence for some children. Children with respiratory problems such as asthma could be absent either because the pollution provokes an attack, or because parents keep the child inside at home in order to avoid pollution exposure. While absences are of interest in their own right, epidemiologists focus on absences as a possibly more sensitive proxy for health status than alternative measures such as emergency room visits or hospital admissions. Children are particularly sensitive to pollution given their small size, high metabolic rates, and developing systems, and it is important to have measures of the effects of pollution that capture its impact on the ability of children to perform their daily activities.

This paper estimates the causal effect of air pollution on elementary and middle school absences using unique administrative data on schooling attendance in 39 of the largest school districts in the state of Texas. These data are merged with information about air quality from monitors maintained by the Environmental Protection Agency (EPA). Although a few other studies examine the relationship between various pollutants and school attendance, they may not have identified causal effects because of the influences of a number of potentially confounding factors.

Absenteeism is affected by parental commitment to schooling, the opportunity cost of a child remaining at home, school policies, and of course health. Given extensive residential sorting based not only on wealth but also on preferences for school quality, environmental quality, peer characteristics, etc., geographic differences in pollution levels are correlated with family characteristics that may in turn be related to absenteeism. For example, pollution tends to be higher where people are poorer, and poor children may be more likely to be absent. In addition, seasonal differences in the prevalence of colds and flu as well as in pollution levels raise doubts about the use of seasonal variation in pollution levels to identify effects on absenteeism.

In order to overcome these problems and to try to identify causal effects, we adopt a difference-in difference-in differences (DDD) strategy, in which we hold characteristics of schools, years, attendance periods and interactions of these variables constant, and focus on variations in pollution within school-year-period cells. Hence, our models control for all fixed characteristics of schools, years, and periods as well as for characteristics of schools in particular years (such as unusual variations in the student body); characteristics of schools in particular attendance periods (such as higher rates of illness in winter); and characteristics of particular years and attendance periods (such as unusual weather patterns affecting multiple school districts). In addition, we also control for precipitation and temperature, two factors that affect air pollution levels and may also affect student health directly.

The evidence supports the contention that pollution affects school attendance. High levels of carbon monoxide (CO), including levels below the regulatory threshold set by the Environmental Protection Agency, increase absenteeism. There is also some, although less certain, evidence that the level of particulate matter (PM) has impacts on attendance. The findings for CO are robust to alternative specifications. For example, our results suggest that reductions in the number of days with high CO levels between 1986 and 2001 in El Paso, an area in Texas with particularly high CO levels, reduced absences by 0.8 percentage points, indicating a significant effect on child health as proxied by school attendance.

The rest of the paper is laid out as follows. Section 2 gives some background information about the health effects of pollution. Section 3 describes our data. Section 4 describes the methodology. Section 5 presents results, and section 6 concludes.

II. Background

A. Pollution and Health

Attention to air quality by the EPA is focused on six primary, or "criteria," air pollutants: ozone, carbon monoxide, nitrogen dioxide, sulfur dioxide, particulate matter and lead. We focus on three of the pollutants (ozone, CO, and PM) that are most commonly tracked by air quality monitors. We dropped nitrogen dioxide because there is no federal standard for these emissions, and nitrogen dioxide is one of several nitrogen oxides that are key components of ozone formation, which we do examine. Data on lead were not available for this study, and airborne lead levels during the 1990s tended to be quite low.[1] Although data on sulfur dioxide (SO_2) were available, SO_2 levels are low enough now that it is not a primary concern. In addition, many of the SO_2 monitors have been removed over time.

Although pollution levels in the U.S. are dropping, high levels remain in many localities. The EPA estimates that 160 million tons of pollution are emitted into the air each year in the U.S. and that in 2002 146 million people were exposed to air that was considered unhealthy at times (EPA, 2003). Many more people are routinely exposed to levels that fall below EPA thresholds but might still cause adverse health effects, as there remains uncertainty over the appropriate threshold levels.[2]

Relatively little is known about the exact mechanisms underlying any health effects, although most pollutants affect the respiratory and cardiovascular systems. Even less is known about the effects of pollutants on children, although it is thought that children are more susceptible to the effects of pollution than adults because their bodies are still developing and because they have higher metabolic rates. For example, a child exposed to the same air pollution source as an adult would breathe in proportionately more air, and suffer proportionately greater exposure. Children also typically spend more time outdoors than adults on average, increasing their total exposure.

Most studies exploring correlations between air pollution and health focus on particulate matter, which is a catch-all term for pollution particles that come from many different sources

[1] See Reyes (2003) for a study of the effects of removing lead from gasoline.

[2] The 24 hour Air Quality Standard (AQS) for PM_{10} is 150 $\mu g/m^3$. The 8 hour AQS for ozone is .08 parts per million (80 ppb). The AQS for CO is 9 ppm.

and can be of different sizes and compositions. Because only small particles can be inhaled into the lungs, the standard is to look at particulate matter less than 10 micrograms per cubic meter of air, or $\mu g/m^3$, in aerodynamic diameter (PM_{10}).[3]

PM has been shown to aggravate and increase susceptibility to respiratory and cardiovascular problems, including asthma. Children and people with existing health conditions are most affected. The leading theory about why PM affects health is that it provokes an immune system response. If, however, immune responses take time to develop, it may be difficult to detect any effect of short-term movements in PM on health outcomes (Dockery, 1993, Hansen and Selte, 1999, EPA 2004, and Samet, 2000, Seaton, et al. 1995).

Ozone (O_3) is a secondary air pollutant formed by nitrogen oxides and volatile organic compounds coming from exhaust, combustion, chemical solvents and natural sources in the presence of heat and sunlight. Ozone has been associated with many respiratory problems and is known to aggravate asthma seriously. Levels rise with the temperature, peaking on hot summer afternoons. There is a considerable amount of within-day variation in ozone levels as the temperature and sunlight changes, with higher levels occurring during the hours people are most likely to be outside. Children who play outside are especially susceptible to ozone. Ozone poses less of a threat indoors, since it quickly reacts with surfaces and becomes harmless. After rising for some time, ozone levels have dropped dramatically in recent years, reaching their 1980 level in 2003 (EPA 2003, and Lippmann, 1992).

---

[3] Fine-particulate air pollution, which includes particles less than 2.5 $\mu g/m^3$ ($PM_{2.5}$), is considered by many experts to be a more effective measure of harmful pollutants because smaller particles are more likely to be made up of toxic materials (such as sulfate and nitrate particles left over from fossil fuel combustion). However, most regulatory agencies only began collecting information on $PM_{2.5}$ in the past few years, so data on $PM_{2.5}$ was not available during the period examined in this paper.

Carbon monoxide is emitted from incomplete combustions occurring in fires, internal combustion engines, appliances, and tobacco smoke. Cars account for as much as 90 percent of CO in some urban areas. CO impairs the transport of oxygen in the body, leading to cardiovascular and respiratory problems. People with pre-existing cardiovascular or respiratory problems appear to be most susceptible to exposure. Levels are highest during cold weather (Lippmann, 1992 and EPA 2004).

The large clinical literature regarding the health effects of pollution has focused primarily on showing associations between air pollution and adverse health outcomes among adults. Nonetheless, because the health of adults reflects a lifetime of exposures in various locations, exposure in their current area of residence may not be a good measure of the effects of pollution.

A smaller literature examines health effects among infants and children, who are more likely to have lived in their current location since birth. This literature has shown associations between air pollution and infant mortality, preterm birth, and low birth weight. However, causal interpretations of these estimates may still be confounded by the fact that pollution tends to be higher in poor and minority areas. Hence, one might expect people in high pollution areas to have worse health for other reasons.

In two important and innovative works, Chay and Greenstone (2003a, b) used changes in regulation to identify pollution's effects on infant mortality. They argue that the 1970 and 1977 Clean Air Acts caused exogenous changes in pollution levels and that the changes varied across areas. These changes can be used to examine pollution's effects on housing markets and infant mortality. They find that a $1\mu g/m^3$ reduction in total suspended particulates (a common measure of overall pollution at that time) resulted in 5-8 fewer deaths infant per 100,000 live births.

Currie and Neidell (2005) examine the effects of air pollution on infant deaths in more recent data. They use individual-level data and within-zip code variation in pollution over time to identify the effects of pollution. They include zip code fixed effects to account for omitted characteristics like ground water pollution and socioeconomic status, and find that reductions in two pollutants – CO and PM_{10} – in the 1990s saved over 1,000 infant lives in California.

Pollution has also been shown to have negative health effects on morbidity in addition to increasing mortality. In a recent study, Neidell (2004) uses within-zip code variation in pollution levels to show that air pollution affects child hospitalizations for asthma. In particular, he finds that if CO levels had been at their 1992 levels in 1998, hospital admissions for asthma would have been 5 to 14 percent greater among children 1 to 18. This is one of the only preceding studies to establish a causal relationship between current levels of pollution and child health.

B. Pollution and Absenteeism

Several epidemiological studies have examined the link between air pollution and absences. In these studies, absences are viewed as a proxy for child health that is more sensitive to pollution induced diseases than hospital related measures. There may be a great deal of illness that is not severe enough to send a child to a hospital, and absence data offers a window on these illnesses. Moreover, there is a long tradition of using absence from school to define disability among children.

Of course pollution is not the only reason for school absences, making it imperative to account for other causes that might confound the estimates. Most absences are due to illness and are attributable either to respiratory infections or to gastroenteritis (Gilliland et al., 2001). However, given the many other factors that can cause absences, it is clearly important to control adequately for a wide range of variables. For example, low socioeconomic status might be correlated both with high exposure to pollution and with high absence rates. Cold, wet weather will contribute to both higher illness and lower pollution levels.

A key difference between economic models and epidemiological models of the effects of pollution is that economists anticipate that parents can respond to potential air pollution through locational choices. Parents can avoid air pollution by moving to neighborhoods with cleaner air, making neighborhood choice an element to consider in modeling the impacts of pollution. Parents and schools can also respond to pollution by keeping children indoors when the air is particularly unhealthful. The decision to keep a child home from school directly adds to absences, and school decisions to keep children indoors may indirectly increase absenteeism if recess indoors contributes to the spread of colds and flu.

We cannot distinguish between absences caused by the direct health effects of pollution and absences caused by avoidance behavior. Nonetheless, to the extent that children miss school in order to avoid pollution, they and their parents still incur a cost, and it is useful capture this type of avoidance behavior as well as the direct effect of pollution on illness in assessing the total costs of pollution.

The recent literature examining the link between absenteeism and pollution is reviewed in Figure 1.[4] Most studies focus on associations between pollution and absences rather than on the identification of causal linkages and control for omitted variables in only a limited way.

One of the most convincing studies is by Ransom and Pope (1992) who examine the impact of PM_{10} on absenteeism in the Utah Valley. Their data encompass a period from August 1986 to September 1987 when a steel mill, the major polluter in the valley, shut down and

[4] One of the first studies of this topic is by Ferris (1970). He studies approximately 700 students from seven schools in Berlin, New Hampshire. He found no difference in the mean levels of absences between schools.

pollution levels fell dramatically. That is, much of the variation in pollution over their sample was due to an event that families had little control over and were unlikely to have responded to in the very short run. They find that absences were 54 to 77 percent higher when PM_{10} increased from 50 to over 100 ug/m^3. These estimates imply that about one percent of students in their sample were absent each day as a result of particulate pollution exposure. In addition, they find that the effect of high PM_{10} levels on absenteeism persisted for 3 to 4 weeks. This persistent effect may complicate attempts to measure responses to short term fluctuations in this pollutant. One important concern with this analysis is the possibility that the job loss following the steel mill closure directly affect absences by reducing the cost of taking care of a child home from school or adversely affecting families and children in ways that reduced school attendance.

A second notable study is by Gilliland et al. (2001) who use data from the Children's Health Study to investigate the effects of pollution on absences due to a number of causes, including respiratory illness. The authors monitored a cohort of 2,081 4th grade children in 12 southern Californian communities from January through June 1996. They focus on the relationship between absence rates and within-community deviations in pollution, and allow for lagged effects of pollution. They find that a 20 parts per billion (ppb) increase in O_3 was associated with a 63 percent increase in absences for illness, and an 83 percent increase in absences for respiratory illnesses. Absences reached a peak five days after exposure.

They also find that, while average PM_{10} levels were associated with higher absence rates, daily increases were associated only with non-illness-related absences. They find this result puzzling, but it could reflect insufficient controls for differences in family background or other variables that are correlated with higher pollution levels. Another possibility is that small cell sizes make the results sensitive to outliers.

The other studies included in Figure 1 have weaker designs. However, Chen et al. (2000) is notable because it is the only previous study to have examined CO. Although CO is known to be dangerous (exposure to high levels is fatal), it has been largely ignored in the epidemiological literature on the effects of pollution and health. Chen et al. find that an increase of 1 ppm in CO increased absence rates by almost 4 percent. Like Gilliland et al. (2001) they also find a curious negative correlation between PM_{10} levels and absences.

Our study is the first to use panel data methods in a large sample in order to estimate the causal effect of air pollution on absenteeism. As we describe in detail below, we try to identify the causal effects of carbon monoxide, ozone, and particulate matter using school differences in the year-to-year variation in pollutant and absence levels within attendance periods. Because we control for seasonal patterns in attendance at each school, idiosyncratic factors affecting a particular school and year, and special factors affecting all schools in a particular year and attendance period, the remaining pollutant differences provide a plausibly exogenous source of variation with which to identify the causal effects of the respective pollutants.

III. Data

Our analysis centers on Texas, which as a large industrial state is a major producer of pollution. In 1998, Texas ranked first in the nation in emissions of nitrogen oxides and volatile organic compounds (the two components of ozone), and second in emissions of CO and PM_{10} (EPA 1998). According to the American Lung Association, the Houston and the Dallas-Fort Worth metropolitan statistical areas rank fifth and tenth, respectively, in listings of areas with the worst ozone pollution in the country. With high and variable levels of pollution, Texas provides a good environment in which to examine the effects of air pollution on school attendance.

This paper uses panel data for schools within 10 miles of pollution monitors in 39 of the largest school districts in Texas –accounting for roughly 40 percent of the students in grades one through eight in the Texas public school system – for the academic years 1996 through 2001. There are 1,512 schools with pollution information in this sample. Two different sources of data are used, one for pollution and one for attendance. Using the latitude and longitude of each school, pollution data is matched with schools. The absence data are only available in aggregated six week attendance period blocks, so the hourly pollution data are also aggregated into six week blocks based on school-specific dates of the attendance periods.[5]

The school data in this paper comes from the Texas Schools Project of the University of Texas, Dallas. The database combines data from a number of different sources and contains student level information for all public school students and teachers in Texas, along with information about the schools themselves. In addition to information about absenteeism, we also use data on gender, family income and ethnicity.[6] Because all students in a school are exposed to the same pollutants, we aggregate the data by school, attendance period, and year to produce the school average absentee rates and demographic characteristics for each six week block in each year. In preliminary work we examined the possibility that pollution effects differed by ethnicity or family income but found no evidence of significant differences. All regressions are

---

[5] Since the school calendars vary from district to district, each district was contacted in order to establish beginning and ending dates for each 6 week period in each year.

[6] By using school absences we are necessarily limited to days when school is in session and therefore omit a large share of the summer period. Although ozone is highly correlated with temperature, it typically peaks in Texas in late August and September when children have returned to school, so we still capture the time of year when ozone levels approach or exceed Air Quality Standards, as Table 1 indicates.

weighted by the number of student observations in the cell. The final sample includes over 12 million student by attendance period observations.

The pollution data come from the Texas Commission on Environmental Quality (TCEQ), formerly the Texas Natural Resource Conservation Commission (TRNCC). Individual monitors for each pollutant are set up all over the state and take hourly readings of the pollution levels at each location.[7] In order to allow for non-linear effects of pollution we measure pollution relative to EPA thresholds and allocate each day into one of five categories separately for each pollutant: 0-25%, 25-50%, 50-75%, 75-100% and greater than 100% of the relevant threshold. We then calculate the shares of days in each category for all six-week attendance periods.[8]

Table 1 provides some descriptive statistics for absences and pollutant levels. The statistics are calculated using data aggregated to the state, year, and attendance period level and are weighted by cell sizes. There are 21,791 cells. The first panel of the table reports mean absence rates and pollution levels by attendance period, means for 1995 and 2000, and maximum values for those two years.

---

[7] Over the years of this analysis, the numbers of monitors for each pollutant changed as some were taken on and off-line. In our sample, the number of O_3 monitors increased from 29 to 41 and the number CO monitors ranged between 20 and 24 between the 1996 and 2001 school years. The number of PM10 monitors actually fell from 38 to 26 between the 1996 and 1998 school years before rising to 31 in the 2001 school year. Changes in the sample composition over time due to changes in the number of monitors are one of the things that will be controlled by the inclusion of school-by-year fixed effects.

[8] Attendance periods with fewer than 25 days of CO and ozone data and five days of PM data are excluded from the sample. The PM threshold is far lower because, as we discuss below, PM readings are recorded roughly every sixth day in many locales.

Absences and pollution levels vary systematically across attendance periods, with both absences and the number of days above 75% of the CO threshold reaching their highest level in the 4th period. Ozone displays the reverse pattern and is at its highest level at the beginning of the school year when absences are at their lowest. There are very few cells with days above the PM threshold but the number of days with PM above 50% of the threshold is highest in attendance periods 3 and 4.

A comparison of the averages for 1995 and 2000 shows that CO and PM levels declined over the study period, as did absences. The largest decline in pollution was in CO, which fell by a substantial 42%. O_3 levels actually increased from the previous year in three of the six years.

Our exploration of the data indicated that O_3 consistently moves in the same direction as temperature, so that differences in ozone pollution from year-to-year are primarily based on changes in average temperature. This highlights the importance of controlling for temperature in the analysis. Precipitation is also like to be important. On the one hand, it cleanses pollutants from the air. On the other hand, rainy weather may be associated with increased absence for illness regardless of the level of pollution.

Table 1 shows that on average, the number of days with high pollution levels is quite small. However, averages mask the fact that while students in most schools in Texas face pollution levels well below EPA standards, a few schools have consistently high pollution levels. The last two columns in the first panel of Table 1 show the maximum values for absences, and for the percentage of days in various categories. For example, in one cell in 1995, 12.8 percent of the school days had pollution levels between 75% and 100% of the EPA threshold for CO, while another cell had 16.67% of days between 75% and 100% of the threshold for PM.

The second panel of Table 1 shows the number of cells with fractions of days in various absence and high pollution categories. While most cells had relatively low levels of pollution, there were at least several hundred cells with high CO and PM levels, and many more with high ozone levels. The high CO cells were located in El Paso, Laredo, and Houston, while high PM cells were located in Galveston and Hidalgo counties.

Table 2 shows the share of child-level observations in each absence or pollution category cell, by the child's race and whether or not the child was receiving a subsidized school lunch.[9] Blacks and Hispanics have somewhat more absences on average than non-Hispanic whites, and school lunch children have more absences than those who do not participate in the program. Turning to the pollution measures, Hispanic children are (perhaps surprisingly) more likely to be exposed to high CO and PM days than either blacks or whites. However, they are less likely to be exposed to high ozone days. Similarly, low income (as measured by eligibility for school lunch) children are more likely to be exposed to CO and PM and somewhat less likely to be exposed to high values of ozone.

IV. Methods

A number of factors complicate efforts to identify the causal effect of pollutants on absenteeism. Both variables that determine pollution levels and those that determine the distribution of families among communities can directly affect absenteeism, making it necessary to account for a number of potentially confounding influences. For instance, if poor or minority families are more likely to live in heavily polluted areas and also have higher absence rates due

[9] Families with incomes less than 130 percent of the federal poverty line are eligible for free school lunches, while families with incomes less than 185 percent of the poverty line are eligible for reduced price meals. This is the best measure of household economic status available in these administrative data.

to factors other than air pollution, then we may obtain spuriously large positive effects of pollution on absences. Given limited information on both family background and pollution determinants, it is highly unlikely that OLS regression models attempting to control for these other influences would produce valid causal estimates. Fortunately, the panel data make it possible to account for unobserved differences in schools, seasons, and years, thereby removing the influences of key confounding factors.

Equation (1) models the absence rate in school s on day d in attendance period p and year y as a function of pollutant levels (P_i is a vector for pollutant i), family and student demographic characteristics X, weather W, a vector G containing the proportions of students in all grades but one, and an error. Because the effects of high pollution days may persist beyond the current day and affect health with a lag, the equation includes the pollutant levels for the current day, previous day, and two days prior. (For expositional ease, the model arbitrarily limits the length of lags to two days, but, given our estimation approach, this is not crucial). In order to allow for nonlinear pollutant effects and to examine the link between federal threshold levels and absenteeism, we include a set of dummy variables for each pollutant that indicates whether the maximum level for the day lies between 25-50 percent of the threshold, 50-75 percent of the threshold, 75-100 percent of the threshold, or is greater than 100 percent of the threshold (0-25 percent is the omitted category). The combination of the lag structure and nonlinear parameterization provides a flexible specification of pollution effects that would capture increasing absenteeism during a series of high pollution days in addition to nonlinearities and time delayed effects. Because all students in a school receive identical pollution treatments, we aggregate over students.[10]

$$(1)\quad A_{sdpy} = \sum_{i=1}^{4} P_{sdpyi}\delta_i + \sum_{i=1}^{4} P_{sdp-1yi}\gamma_i + \sum_{i=1}^{4} P_{sd-2pyi}\lambda_i + X_{sdpy}\beta + W_{sdpy}\theta + G_{sdpy}\varphi + \varepsilon_{sdpy}$$

A key factor influencing our estimation strategy is the aggregation of absence information during the school year into six separate six-week blocks. This leads us to aggregate the pollution information into six-week blocks as well. Equation (2) aggregates equation (1) by school, attendance period, and year and also decomposes the error into a series of terms that illustrate the key unobserved influences accounted for by the panel data methods. The aggregation incorporates the almost perfect collinearity between period aggregates of pollution distributions for current, one day, and two day lagged pollution values and makes explicit that the aggregate nature of the attendance data precludes identification of the precise timing of pollutant effects. Consequently the pollutant coefficients provide the cumulative effects of a pollution event on absences regardless of time after exposure.

$$(2)\quad \overline{A}_{spy} = \sum_{i=1}^{4} \overline{P}_{spyi}(\delta_i + \gamma_i + \lambda_i) + \overline{X}_{spy}\beta + \overline{W}\phi + \overline{G}\varphi + \rho_s + \pi_p + \tau_y + \mu_{sp} + \phi_{sy} + \eta_{yp} + \omega_{spy}$$

We control for confounding factors with information on observable characteristics and by a set of fixed effects made possible by the multiple schools, attendance periods, and years of

[10] As will be clear, this aggregation also makes the estimation computationally tractable. While it is possible to formulate the model with differential effects across individuals, our preliminary investigations found no systematic differences in pollution impacts by race or income, leading us to restrict our attention to the models with common school effects. Remember, however, that the levels of ambient pollution differ by ethnicity because of sorting among neighborhoods and schools, so that the common marginal effect of pollution translates into differential total impacts.

data. In terms of observables, X includes the shares of students who are Asian, Hispanic, and black (whites are the omitted category), the share of students eligible for a subsidized lunch, and the share of females; and W includes measures of temperature and precipitation. In terms of unobservables, the error components include school, attendance period, and year fixed effects along with all pair wise interactions and a random error ω.

The inclusion of X and the school fixed effects (ρ_s) accounts for persistent between school differences in average student characteristics that might be correlated with absences. The grade shares G account for any systematic differences in absenteeism related to age. Year dummies (τ_y) control for intertemporal patterns in attendance rates, and attendance period dummies (π_p) control for seasonal effects. The pair wise interactions among school, year, and period control for many other factors that might be related to absences. For example, school*year effects (φ_{sy}) account for changes over time in student populations or local variations in annual weather, while attendance period*school effects (μ_{sp}) help to account for distinctive seasonal patterns within schools (for example, schools with high immigrant enrollments might routinely experience high absences surrounding the Christmas vacation period). Finally, year*period effects (η_{yp}) account for unusual seasonal effects affecting all districts, such as an unusually bad flu season.

Given all of these controls, the causal effect of pollution is identified by variation at the year-school-period level, illustrated by the following example. Consider looking at a particular Houston elementary school in 2000 and comparing attendance rates in the highest and lowest pollution periods within that year. Because we are looking within a school and within a year, the "effect" of pollution would be identified by the difference in pollution between the two periods. Alternatively, we could compare attendance rates in a single season in high pollution and low pollution years. In this case the “effect” of pollution would be identified by the difference in pollution between the two years.

Attendance might nonetheless vary across seasons or within season across years for reasons independent of the effects of pollution. To control for this possibility, we could make use of the multiple years of data and control for school average pollution levels for each season and for each year. In this case the “effect” of pollution would be identified by deviations from a school’s average pollution levels for each season and year. In short, one could estimate a difference in differences model.

We could go further to guard against other unobserved differences by attendance period and year that could contaminate the estimates. The availability of data across schools would allow us to control for average period-by-year effects across all schools. The inclusion of school-by-year, school-by-attendance period, and attendance period-by-year fixed effects (along with other measurable variables that differ by school, attendance period, and year) would thus control for myriad potentially confounding factors, and this difference-in-difference-in-differences (DDD) model is precisely that used in the analysis.

The DDD estimates will generate consistent estimates of the effects of pollution as long as there are no other factors which vary with pollution at the school-year-period level, and which also affect absences. An example of something which would violate this identification assumption would be a natural disaster (such as a forest fire) which increased air pollution levels and also resulted in increased absences in a particular school, year, and period. While we cannot rule this possibility out, we are unaware of any such incidents during our sample period.

V. Results

This section reports the results for a series of specifications based on equation (2) along with variants designed to understand the robustness of the estimation. The alternative estimates differ along a number of dimensions including criteria for school inclusion in the sample, the specific set of fixed effects and interactions, and the included pollutants and their parameterization. All specifications include period-by-year fixed effects, the grade share variables, demographic characteristics, and weather variables and are weighted by the number of students in each school-period-year cell. Because pollutant levels vary by pollution monitor, attendance period, and year, robust standard errors clustered at that level are used to calculate t statistics. Despite the large numbers of students and over twenty thousand school-by-period-by-year combinations, the number of monitor-period combinations is around 630.

Table 3 reports results for a series of specifications estimated over a sample of schools within 10 miles of a pollution monitor. The first column shows estimates from a specification without school, school-by-attendance period, or school-by-year fixed effects, and the subsequent columns illustrate changes to the estimates as these fixed effects are successively added.

Table 3 reveals substantial variation in both the estimated relationship between the rate of absenteeism and the individual pollutants and the sensitivity of the coefficients to the inclusion of additional covariates and fixed effects. The evidence of a negative pollutant effect is clearly strongest for CO. All coefficients on the shares of days 75-100 percent of the threshold and greater than 100 percent of the threshold are positive, and once school-by-attendance period fixed effects are included they are ordered in the expected direction and significant (Column 3). The addition of school-by-year fixed effects increases the magnitude and significance of the coefficients (Column 5), providing strong support for the existence of a causal effect given the much more restricted variation used to identify the estimates in the presence of school-by-year fixed effects.

Finally, the much smaller and statistically insignificant coefficients on the shares of days less than 50 percent of the threshold along with the much larger effect for days greater than 100 percent of the threshold provide support for a nonlinear relationship between CO pollution and the probability of being absent. An additional day that the CO level exceeds the threshold increases absenteeism by almost 9 percentage points, while an additional day between 75 and 100 percent of the threshold increases absenteeism by 5 percentage points.[11] Although such high levels of CO are not commonplace during this period, Table 1 did show that at least 10 percent of the days exceeded 75 percent of the CO threshold in 66 school-by-attendance period-by-year cells. Thus the estimates imply increases in the period average absenteeism rate of roughly one half a percentage point in these cells, a significant effect but much smaller than that in Chen et al (2000) which did not control as thoroughly for potential confounders.

In contrast, the proportion of days ozone exceeds the EPA threshold has a significant positive effect on absences in specifications without the fixed effects. However, once the fixed effects are added the coefficient moves closer to zero and is statistically insignificant (Column 5). It appears that differences between communities and across seasons generate the positive relationship between absences and the share of days that ozone exceeds the threshold in the models without fixed effects.

[11] Technically, the estimates say that if the share of high pollution days in the attendance period increased to 100 percent, then absences over the attendance period would increase by either 9 or 5 percentage points. We can assume that this relationship also holds for daily attendance, as long as daily attendance is affected by daily pollution levels (rather than by cumulative pollution levels over several days, for example).

We also find a robust and significant increase in absences associated with PM levels of 50-75% of the EPA threshold, but no significant effect from days over 75% of the threshold. The relative rarity of higher PM level days and measurement error introduced by the infrequent reporting of PM levels in most monitors raises the possibility that measurement error for the PM variable conceals the adverse effects of exposure to high levels. As opposed to ozone and CO for which we have daily information on maximum levels, PM measurements for most monitors are typically available one of every six days.[12] The very small number of days in which PM levels exceed even 75 percent of the threshold (see Table 1) indicates that the true variances of the shares of days 75-100 percent and greater than 100 percent of the threshold are quite small, compounding the difficulty of generating a precise estimate. Given the multiple fixed effects in the empirical models, it is not surprising that these coefficients are estimated imprecisely. If this is the correct interpretation of our results, then they suggest that PM levels as low as 50 percent of the EPA threshold may have deleterious health effects, an important result in light of continuing debate about whether EPA should lower the thresholds for PM_{10} .[13]

In order to investigate the potential effects of measurement error and multicollinearity among the pollutants, we estimated a series of additional models. Table 4 replicates Table 3 for the sample of schools within 5 miles of monitors for all three pollutants, under the assumption that readings provide more accurate measures of exposure the closer a school is to a monitor. This restriction reduces the sample size by roughly two-thirds, but the pattern of estimates is

[12] A few monitors report daily or close to daily readings, but the limited data introduce substantial measurement error for all of the PM indicator variables.

[13] In Sept. 2006, the EPA revised the standard for $PM_{2.5}$ from 65 to 35 $\mu g/m^3$. However, EPA did not change the standard for PM_{10} because of lack of scientific evidence linking PM_{10} below the threshold to health consequences.

quite similar to that observed in Table 3. In particular, CO levels between 75% and 100% and above 100% of the EPA threshold remain associated with significantly higher absence rates.

One notable difference between Table 3 and Table 4 is that we find a significant effect of ozone levels above the EPA threshold in regressions that include a single dummy for whether the threshold is exceeded or not (see column 6 of Table 4). The finding raises the possibility that high levels of ozone increase absenteeism, but the result is not robust to including the indicators for the proportion of days with lower pollution levels.

Table 5 reports results from two additional full fixed effects models designed to illuminate the sensitivity of the estimates to both specification and the structure of the pollution data. The first column shows estimates for single pollutant models estimated over the full sample of schools. That is, the column shows the results of three separate regressions, one for each pollutant. The pattern of results for all pollutants is quite similar to that found in Table 3, suggesting multicollinearity is not hampering our analysis.

Column 2 of Table 5 reports results from the full fixed effect specification (Column 5 of Table 3) regressed on a sample in which a school's pollutant levels for CO and ozone are based only on information for days in which the PM monitor closest to the school has a reading for PM. Again the CO results are quite similar to those reported in Table 3, probably because the monitors that report CO levels above the EPA threshold appear to collect daily PM information so that the data restriction has little effect on the shares of reported days above the threshold.

VI. Conclusion

Obtaining convincing estimates of the causal impact of air pollution on health is difficult. Two major obstacles include the presence of confounding brought about through residential sorting and the lack of health measures that capture the range of morbidities purportedly related to pollution. Through merging administrative school-level panel data with pollution monitor data, we confront these issues by holding school, year, and attendance period characteristics constant to control for many confounding factors and by linking air quality directly with school absences, a sensitive measure of children's health and morbidity.

Our major finding is a significant and robust effect of CO on school absences, both when CO exceeds air quality standards (AQS) and when CO is 75-100% of AQS. Although the number of days where CO approaches these levels is quite low over the time period we study, this has not always been the case. For example, in 1986 in El Paso, an area in Texas with particularly high CO levels, there were 16 days that CO exceeded AQS and 19 days where CO was between 75 and 100% of the threshold. The respective numbers in 2001, the last year for which we have attendance data available, are 1 and 6. Based on our econometric estimates, this decrease in high CO days decreased absences in 2000-01 school year by 0.8 percentage points, a significant change compared to the baseline absence rate in all of Texas of 3.58%. Since the improvements in air quality over time are largely due to air quality regulations (Henderson, 1996, Chay and Greenstone, 2003b), this suggests air quality regulations are having a positive impact on children's health.

These effects become even more pronounced if we consider that the effects of pollution concentrate on vulnerable segments of the populations, such as asthmatics. Although we cannot identify asthmatic children in our data, we provide the following back-of-the-envelope calculation to assess the possible magnitude of effects for asthmatics in El Paso. The asthma prevalence rate in 2002 for school age children in the U.S. is 14.0 percent and, conditional on being absent, the probability the child has asthma is 28.3%.[14] Based on the mean absence rate in

[14] Based on authors' calculations from the 2002 National Health Interview Survey.

our sample of 3.58%, this implies an absence rate of 7.2% among asthmatics. If all of the 0.8 percentage point reduction in absences in El Paso occurred among asthmatics, this would imply that the absence rate among asthmatics was reduced by roughly 44%.

This finding for CO, combined with growing evidence from other recent studies documenting a robust effect of CO on different populations and outcomes (Currie and Neidell, 2005, and Neidell, 2004), speaks to the continuing debate over regulations concerning automobile emissions (Barnes and Eilperin, 2007). CO is a product of incomplete combustion of fuels, making automobiles a major source of CO in urban areas, contributing as much as 90% of CO levels. There are three primary approaches for reducing CO emissions: technological innovations in fuel combustion (such as catalytic converters), development of alternative fuels that do not emit CO (such as ethanol, biodiesel, and fuel cells), and reductions in vehicle miles traveled. Despite dramatic improvements in emissions technology over the past several decades, increases in vehicle miles traveled offset some of these improvements. Unless significant improvements in CO emissions can be made, further efforts at reducing CO must therefore also focus directly on consumer behavior, via mechanisms such as congestion pricing and emissions taxes. Whichever approach is taken, it will have the additional benefit of reducing carbon dioxide, the primary greenhouse gas linked with climate change.

References

Action for Healthy Kids, 2004. "Students' Poor Nutrition and Inactivity Comes With Heavy Academic and Financial Costs to Schools," Sept. 2004. http://actionforhealthykids.org/filelib/pr/PoorNutritionIncursCosts-PressRelease.pdf

Barnes, Robert and Juliet Eilperin. "High Court Faults EPA Inaction on Emissions," Washington Post, April 3, 2007, A01.

Bridges to Sustainability and Mickey Leland National Urban Air Toxics Research Center. 2003. "Final Report: Assessment of Information Needs for Air Pollution Health Effects Research in Houston, Texas." Feb. 17, 2003.

Chay, Kenneth and Michael Greenstone 2003a. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper #10053.

Chay, Kenneth and Michael Greenstone 2003b. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118(3): 1121-1167.

Chen, Lei et al. 2000. "Elementary School Absenteeism and Air Pollution." *Inhalation Toxicology*. 12: 997-1016

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

Dixon, Jane. 2002. "Kids Need Clean Air: Air Pollution and Children's Health." *Family & Community Health.* 24(4): 9-26.

Dockery, Douglas et al. 1993. "An Association Between Air Pollution and Mortality in Six U.S. Cities." *The New England Journal of Medicine.* 329 (24): 1753-1759

Ferris, Benjamin. 1970. "Effects of Air Pollution on School Absences and Differences in Lung Function in First and Second Graders in Berlin, New Hampshire, January 1966 to June 1967." *American Review of Respiratory Disease*. 102: 591-606.

Gilliland, Frank et al. 2001. "The Effects of Ambient Air Pollution on School Absenteeism Due to Respiratory Illness." *Epidemiology*. 12: 43-54.

Grossman, Michael and Robert Kaestner. 1997. "The Effects of Education on Health," in The Social Benefits of Education, Jere Behrman and Nancy Stacy (editors) (Ann Arbor: University of Michigan Press).

Hansen, Anett and Selte, Harald. 2000. "Air Pollution and Sick-Leaves," *Environmental and Resource Economics*. 16: 31-50, Henderson, J. Vernon. 1996. "Effects of Air Quality Reguations," *American Economic Review* 86: 789-813.

Houghton, F et al. 2003. "The Use of Primary/National School Absenteeism as a Proxy Retrospective child Health Status Measure in an Environmental Pollution Investigation." *Public Health*. 117: 417-423.

Linn, William et al. 1996. "Short-Term Air Pollution Exposures and Responses in Los Angeles Area Schoolchildren." *Journal of Exposure Analysis and Environmental Epidemiology* 6 (4):449-472

Lippman, Morton. "Environmental Toxicants." Van Nostrand Reinhold, New York. 1992.

Makino, Kuniyoshi. 2000. "Association of School Absences with Air Pollution in Areas around Arterial Roads." *Journal of Epidemiology*. 10 (5): 292-299.

Neidell, Matthew 2004. "Air Pollution, Health, and Socio-Economic Status: The Effect of Outdoor Air Quality on Childhood Asthma," *Journal of Health Economics* 23(6).

Park, Hyesook et al. 2002. "Association of Air Pollution with School Absenteeism Due to Illness." *Archives of Pediatric & Adolescent Medicine.* 156: 1235- 1239.

Pastor, Manuel et al. 2004. "Reading, Writing and Toxics: Children's Health, Academic Performance, and Environmental Justice in Los Angeles." *Environment and Planning C: Government and Policy.* 22: 271-290.

Ransom, Michael and Pope, C. 1992. Arden. "Elementary School Absences and PM10 Pollution in Utah Valley." *Environmental Research.* 58: 204-219.

Reyes, Jessica. "The Impact of Prenatal Lead Exposure on Infant Health." Working Paper, 2003.

Reyes, Jessica. "Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime." Working Paper, 2002.

Samet et al. 2000. "Fine Particulate Air Pollution and Mortality in 20 U.S. Cities, 1987-1994." *The New England Journal of Medicine*. 343 (24):1742-1749.

Seaton, Anthony et al. "Particulate Air Pollution and Acute Health Effects," The Lancet, v354, Jan. 21, 1995, 176-178.

Sunyer et al. 1997. "Urban Air Pollution and Emergency Admissions for Asthma in Four European Cities: The APHEA Project." *Thorax*. 52: 760-765.

U.S. Environmental Protection Agency. Benefits and Costs of the Clean Air Act 1990-2020: Draft Analytical Plan for EPA's Second Prospective Analysis, Office of Policy Analysis and Review, Washington D.C., 2001.

U.S. Environmental Protection Agency. National Trends in CO Levels. Washington D.C., updated April 2007. http://www.epa.gov/airtrends/carbon.html.

U.S. Environmental Protection Agency. Air Emissions Trends – Continuing Progress Through 2003. Washington D.C., 2004 http://www.epa.gov/airtrends/econ-emissions.html.

U.S. Environmental Protection Agency. Six Principal Pollutants. Washington D.C., 2004. http://www.epa.gov/airtrends/sixpoll.html.

U.S. Environmental Protection Agency. Executive Summary of 2003 National Air Quality and Emissions Trends Report. Washington D.C., 2003. http://www.epa.gov/air/airtrends/aqtrnd03/pdfs/chap1_execsumm.pdf.

U.S. Environmental Protection Agency. 1998 National Air Quality and Emissions Trends Report. Washington D.C., 1998. http://www.epa.gov/air/airtrends/aqtrnd98/.

Authors	Sample	Method	PM_{10}	CO	O_3	NO_X
		Figure 1: Previous Literature				
Ransom and Pope, 1992	Weekly absenteeism data from Provo Utah school district, daily data from one school in Alpine Utah, 1986-1991.	Percent absent for each grade regressed on PM, controlling for day of week, month of school year. Data incorporates period of plant shutdown when pollution fell dramatically.	+	N/A	N/A	N/A
Gilliland et al. 2001	Six months of data from 1996 on school absences, and reasons for absences for schools in 12 CA communities. Hourly pollution measures.	Regress absence rate on within-community deviation in pollution from average levels. Examine effects on illness and non-illness related absences separately.	-	N/A	+	0
Makino, 2000	Two schools in Japan located near arterial roads. Daily attendance/pollution data over 1993-1997.	Regresses daily attendance data on PM and NO_X. Omits winter data due to flu season.	+	N/A	N/A	+
Chen et al. 2000	57 schools in Washoe County NV, 1996-1998.	Absence rate for each grade regressed on pollutants, weather, day of week, month, holiday, time trend.	-	+	+	N/A
Park et al. 2002	One elementary school in South Korea, from March 1996-Dec. 1999.	Daily absences regressed against daily pollutant levels using a Poisson regression.	+	N/A	+	0

Table 1: Distribution of Pollution levels and Absences in School*Grade*Year*Attendance Period Cells

	Average Period 1	Average Period 2	Average Period 3	Average Period 4	Average Period 5	Average Period 6	Average All Periods	Average All-1995	Average All-2000	Max All-1995	Max All-2000
Weighted averages over cells in columns 1 to 9, maximum cell values in columns 10 and 11.											
Percent days absent	2.35	3.21	3.97	4.30	4.04	3.72	3.58	3.77	3.35	15.66	11.18
CO											
percent of days 0-25% threshold	97.06	93.43	86.98	90.26	95.01	98.10	93.46	89.35	97.39		
percent of days 25-50% threshold	2.92	5.87	10.86	8.41	4.61	1.88	5.77	9.12	2.45	61.54	65.10
percent of days 50-75% threshold	0.02	0.51	1.93	1.04	0.33	0.02	0.65	1.30	0.15	20.51	13.95
percent of days 75-100% threshold	0.00	0.14	0.23	0.26	0.05	0.00	0.11	0.23	0.01	12.82	2.56
percent of days > 100% threshold	0.00	0.05	0.00	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00
OZ											
percent of days 0-25% threshold	6.42	12.71	32.58	24.75	7.72	4.25	14.51	11.44	16.63		
percent of days 25-50% threshold	22.38	44.86	59.37	64.90	47.08	31.30	44.78	47.01	44.89	92.30	92.30
percent of days 50-75% threshold	27.75	29.88	7.47	10.02	40.24	46.42	26.78	27.09	25.20	85.00	91.30
percent of days 75-100% threshold	26.87	10.51	0.55	0.32	4.55	17.54	10.20	9.63	10.06	51.28	52.63
percent of days > 100% threshold	16.58	2.04	0.03	0.01	0.41	0.49	3.73	4.83	3.22	52.94	35.90
PM											
percent of days 0-25% threshold	77.88	85.34	86.73	87.34	85.73	83.40	84.34	80.20	86.81		
percent of days 25-50% threshold	21.51	13.75	12.01	11.78	13.33	15.15	14.65	18.58	12.57	83.33	85.70
percent of days 50-75% threshold	0.61	0.61	1.06	0.72	0.69	1.35	0.84	1.07	0.61	33.33	16.67
percent of days 75-100% threshold	0.00	0.23	0.20	0.05	0.11	0.05	0.11	0.12	0.00	16.67	0.00
percent of days > 100% threshold	0.00	0.07	0.00	0.11	0.14	0.05	0.06	0.03	0.01	2.78	7.14

Number of school x grade x attendance period x year cells with percentage of days in various categories						
	0	0-2.5%	2.5-5%	5-7.5%	7.5-10%	>10%
Days absent	26	4475	14970	2204	91	25
CO						
days 75-100% threshold	21253	150	211	61	50	66
days > 100% threshold	21715	42	2	32	0	0
OZ						
days 75-100% threshold	8211	1153	1540	1250	1459	8178
days > 100% threshold	14978	998	1698	827	541	2749
PM						
days 75-100% threshold	21408	0	202	78	24	79
days > 100% threshold	21617	0	86	2	25	61

Note: There are approximately 30 school days and 42 total days in each attendance period. The percent of days absent is calculated using only school days, while the number of days in each pollution category is calculated using the total number of days. There are 21,791 school by year by attendance year cells.

Table 2: Distribution of Pollution levels and Absences by Race/Ethnicity and Income

	Blacks	Hispanics	Whites	School Lunch	No School Lunch
Average proportion days absent	0.0359	0.0359	0.0335	0.0379	0.0309
CO					
proportion of days 25-50% threshold	0.0473	0.0585	0.0423	0.0558	0.0445
proportion of days 50-75% threshold	0.0035	0.0074	0.0035	0.0064	0.0039
proportion of days 75-100% threshold	0.0003	0.0015	0.0005	0.0011	0.0006
proportion of days > 100% threshold	0.0000	0.0002	0.0000	0.0001	0.0000
OZ					
proportion of days 25-50% threshold	0.4395	0.4570	0.4361	0.4514	0.4389
proportion of days 50-75% threshold	0.2641	0.2829	0.2838	0.2765	0.2824
proportion of days 75-100% threshold	0.1126	0.0962	0.1178	0.1010	0.1151
proportion of days > 100% threshold	0.0426	0.0319	0.0468	0.0346	0.0458
PM					
proportion of days 25-50% threshold	0.1380	0.1629	0.1013	0.1573	0.1078
proportion of days 50-75% threshold	0.0053	0.0097	0.0023	0.0084	0.0032
proportion of days 75-100% threshold	0.0001	0.0014	0.0002	0.0011	0.0003
proportion of days > 100% threshold	0.0001	0.0009	0.0002	0.0006	0.0002

Table 3. Estimated Effects of CO, PM10, and Ozone on Average Days Absent.
(schools within 10 miles of pollution monitors).

school fixed effects	n	y	y	y	y	y
school by attendance period fixed effects	n	n	y	n	y	y
school by year fixed effects	n	n	n	y	y	y
Proportion of days CO was						
25-50% EPA threshold	0.0004	-0.0012	-0.0002	-0.0027	-0.0018	
	(0.16)	(0.56)	(0.09)	(1.33)	(1.05)	
50-75% EPA threshold	-0.0177	-0.0056	-0.0069	-0.0068	-0.0110	
	(1.89)	(0.82)	(0.99)	(1.11)	(1.96)	
75-100% EPA threshold	0.0362	0.0125	0.0229	0.0271	0.0500	
	(2.28)	(1.09)	(1.73)	(2.81)	(4.30)	
> 100% EPA threshold	0.0279	0.0172	0.0379	0.0449	0.0891	0.0928
	(1.05)	(1.26)	(2.01)	(1.71)	(4.21)	(3.08)
Proportion of days Ozone was						
25-50% EPA threshold	0.0042	-0.0023	-0.0032	-0.0031	-0.0033	
	(2.20)	(1.90)	(2.72)	(2.58)	(3.10)	
50-75% EPA threshold	0.0022	-0.0080	-0.0042	-0.0090	-0.0031	
	(1.09)	(5.13)	(3.10)	(5.50)	(2.26)	
75-100% EPA threshold	0.0050	-0.0030	0.0002	-0.0042	0.0010	
	(1.54)	(1.69)	(0.15)	(2.01)	(0.59)	
> 100% EPA threshold	0.0093	-0.0070	-0.0055	-0.0060	-0.0007	0.0023
	(2.94)	(3.21)	(2.90)	(2.55)	(0.34)	(1.27)
Proportion of days PM was						
25-50% EPA threshold	-0.0041	0.0012	0.0016	-0.0001	0.0003	
	(4.15)	(1.71)	(2.86)	(0.16)	(0.47)	
50-75% EPA threshold	0.0071	0.0086	0.0060	0.0082	0.0049	
	(1.83)	(4.09)	(2.53)	(3.43)	(2.71)	
75-100% EPA threshold	0.0148	0.0074	0.0009	0.0143	0.0054	
	(2.02)	(1.60)	(0.16)	(2.29)	(1.04)	
> 100% EPA threshold	-0.0005	-0.0085	-0.0063	-0.0145	-0.0112	-0.0089
	(0.05)	(1.47)	(0.87)	(2.05)	(1.74)	(1.34)

Notes: Data aggregated to school, year, attendance period level. Regressions weighted by cell size. 21,791 obs. and 634 monitor, period, year cells. Parentheses show T-statistics based on robust standard errors clustered by monitor*attendance period*year.

Table 4. Effects of CO, PM10 and Ozone on Attendance
(schools within 5 miles of pollution monitors).

school fixed effects	n	y	y	y	y	y
school by attendance period fixed effects	n	n	y	n	y	y
school by year fixed effects	n	n	n	y	y	y
Proportion of days CO was						
25-50% EPA threshold	0.0083	-0.0002	0.0009	-0.0020	0.0000	
	(3.72)	(0.10)	(0.51)	(0.92)	(0.00)	
50-75% EPA threshold	-0.0073	0.0004	-0.0046	0.0029	-0.0002	
	(0.95)	(0.07)	(0.71)	(0.47)	(0.04)	
75-100% EPA threshold	0.0113	0.0044	0.0084	0.0171	0.0306	
	(0.84)	(0.38)	(0.69)	(1.85)	(2.75)	
> 100% EPA threshold	0.0244	0.0096	0.0267	0.0301	0.0642	0.0728
	(1.03)	(0.53)	(1.24)	(1.41)	(2.88)	(2.66)
Proportion of days Ozone was						
25-50% EPA threshold	0.0008	-0.0035	-0.0040	-0.0040	-0.0043	
	(0.42)	(2.54)	(2.54)	(2.98)	(3.05)	
50-75% EPA threshold	0.0013	-0.0086	-0.0036	-0.0087	-0.0018	
	(0.64)	(4.99)	(2.10)	(5.08)	(1.18)	
75-100% EPA threshold	0.0039	-0.0051	-0.0014	-0.0070	-0.0012	
	(1.33)	(2.23)	(0.60)	(2.91)	(0.59)	
> 100% EPA threshold	0.0079	-0.0085	-0.0065	-0.0049	0.0026	0.0059
	(2.24)	(3.33)	(2.20)	(1.84)	(0.93)	(2.40)
Proportion of days PM was						
25-50% EPA threshold	-0.0028	-0.0006	0.0000	-0.0012	-0.0003	
	(2.34)	(0.73)	(0.08)	(1.33)	(0.48)	
50-75% EPA threshold	0.0098	0.0084	0.0064	0.0083	0.0059	
	(3.43)	(3.54)	(2.71)	(3.34)	(3.14)	
75-100% EPA threshold	0.0060	0.0079	0.0041	0.0198	0.0138	
	(0.91)	(1.44)	(0.66)	(3.34)	(2.15)	
> 100% EPA threshold	-0.0077	-0.0090	-0.0057	-0.0147	-0.0118	-0.0042
	(0.59)	(1.11)	(0.68)	(2.06)	(1.80)	(0.60)

Notes: Data aggregated to school, year, attendance period level. Regressions weighted by cell size. 7,613 obs. and 585 monitor, period, year cells. Parentheses show T-statistics based on robust standard errors clustered by monitor*attendance period*year.

Table 5: Additional Models of the Proportion of Days Absent
(Schools within 10 miles of a monitor)

school fixed effects	y	y
school by attendance period fixed effects	y	y
school by year fixed effects	y	y
	Single Pollutant Models	**CO and Ozone measured on same days as PM**
Proportion of days CO was		
25-50% EPA threshold	-0.0020	-0.0015
	(1.16)	(0.82)
50-75% EPA threshold	-0.0095	-0.0096
	(1.65)	(1.58)
75-100% EPA threshold	0.0501	0.0512
	(4.34)	(4.33)
> 100% EPA threshold	0.0973	0.0875
	(4.80)	(4.17)
Proportion of days Ozone was		
25-50% EPA threshold	-0.0034	-0.0032
	(2.90)	(3.08)
50-75% EPA threshold	-0.0032	-0.0032
	(2.34)	(2.30)
75-100% EPA threshold	0.0009	0.0010
	(0.51)	(0.53)
> 100% EPA threshold	-0.0009	-0.0006
	(0.42)	(0.29)
Proportion of days PM was		
25-50% EPA threshold	0.0006	0.0002
	(1.00)	(0.35)
50-75% EPA threshold	0.0055	0.0049
	(3.47)	(2.64)
75-100% EPA threshold	0.0098	0.0058
	(1.61)	(1.05)
> 100% EPA threshold	-0.0109	-0.0138
	(1.52)	(2.12)
observations	21,791	21,321
monitor by period by year combinations	634	628

Notes: data aggregated to school by year by attendance period level; regressions weighted by cell size; absolute value of t statistics based on robust standard errors clustered by monitor*attendance period*year in parentheses.