Summer 8-7-2020

# The Effects of Air Pollution on School Attendance in the United States

Debanhi G. Alrich
*CUNY Hunter College*

The Effects of Air Pollution on School Attendance in the United States

By

Debanhi Alrich

Submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Economics, Hunter College
The City University of New York

Summer 2020

Thesis Sponsor:

| | |
|---|---|
| 1 August, 2020 | Matthew Baker |
| Date | First Reader |
| 1 August, 2020 | Randall K. Filer |
| Date | Second Reader |

## Abstract

This study employs pooled OLS and fixed effect regressions to examine the effects of ambient air pollution on self-reported school attendance in the 423 most populous counties in the United States. Information from the U.S. Census' American Community Survey's annual estimate is compared against county-level data for 14 common air pollutant variables. When making this comparison for the general population, we find statistically significant results for only one pollutant: ozone. We find further significant effects for the presence of ozone when respondents are grouped by race/ethnicity and by poverty ratio, indicating that effects of pollution, like many other social ills, may be borne disproportionately by the poor and people of color.

# Table of contents

# I. Introduction

The connection between ambient (i.e. from the surrounding environment) air pollution and health outcomes is well-established. The World Health Organization (WHO) found that, in 2016, some 91 percent of the world's population was living in areas where air pollution levels exceed guidelines set by the organization in 2005[1]. WHO also found that at least 4.2 million premature deaths worldwide in 2016 were caused by elevated levels of air pollution, particularly in low- and middle-income countries in Southeast Asia and the Pacific Rim. Poor air quality contributes to a range of adverse health conditions, including stroke, heart disease, lung cancer, and chronic and acute respiratory diseases such as asthma (WHO 2018).

Unlike much of the world, the United States has seen a general improvement in air quality. The U.S. Environmental Protection Agency (EPA) sets national ambient air quality standards (NAAQS) for six "criteria pollutants"[2] as mandated by the Clean Air Act. Since standards were last revised in 1990, average national levels for all six categories have declined steadily, and all have remained below threshold levels since 2013. Despite this, areas of concern do exist. So-called "nonattainment areas" have

[1] "2005 WHO Air quality guidelines offer global guidance on thresholds and limits for key air pollutants that pose health risks...The Guidelines apply worldwide and are based on expert evaluation of current scientific evidence for particulate matter (PM); ozone ($O_3$); nitrogen dioxide ($NO_2$); sulfur dioxide ($SO_2$). The WHO Air quality guidelines are currently under revision with an expected publication date in 2020." (WHO 2018)

[2] These "criteria pollutants" are carbon monoxide (CO); lead (Pb); nitrogen dioxide ($NO_2$); ozone ($O_3$) particulate matter (PM); and sulfur dioxide ($SO_2$) (EPA 2020)

been identified by the agency throughout the U.S., mostly in urban areas and especially in Southern California and throughout the Northeast (EPA 2020).

School-aged children likely bear the effects of poor air quality disproportionately. The WHO found that air pollution in certain European cities had a direct impact on issues such as “lung function, childhood infections, the development and severity of allergic diseases (including asthma), childhood cancer and neurobehavioral development” (WHO 2005). A review of the literature on this subject in Bates (2005) concludes that “there is no doubt that relatively low levels of pollution are responsible for increased morbidity” in children, noting that “children may have relatively high ambient pollution exposures, partly because they are physically active out of doors”. Despite this, most inquiries into the effect of air pollution on health focus on adult outcomes.

Less information exists about the social impacts of air pollution on children’s well-being, particularly on educational attainment. One growing area of study is the effect of air quality on school absences, the latter being a key predictor of a child’s eventual educational achievement level. Students with high levels of absenteeism experience lower grades and a higher drop-out rate (Grossman & Kaestner 1997). Frequent absentees have also been shown to underperform in math and reading and exhibit lower educational and social engagement (Gottfried 2014).

This paper seeks to add to the growing body of literature linking poor air quality to increased absenteeism by examining the question at the national level, and over a period of years rather than weeks or months. Other literature has examined local conditions over a narrow band of time (Ransom and Pope 1992; Ransom and Pope

2013) or, at best, studied the issue at the state level (Currie et al. 2009). Also, using self-reported Census data allows for the examination of the effects when controlled for social factors such as Race, ethnicity and income.

# II. Background

This study examines EPA data related to five of the six "criteria pollutants": carbon monoxide (CO), nitrogen dioxide (NO), ozone ($O_3$), particulate matter (PM), and sulfur dioxide ($SO_2$). (The sixth, lead, has been at very low levels for decades to the point where data in the U.S. is all but nonexistent.) Before reviewing the literature examining the effects of these pollutants on school attendance, we will briefly define each and explore their common health effects.

## 2.1 Definition of pollutants

*Carbon Monoxide* (CO) is a colorless and odorless gas originating mostly from internal combustion engines of vehicles. CO is formed "during the incomplete combustion of carbon-containing fuels. While complete combustion leads to the formation of carbon dioxide, most combustion systems involve some fuel-rich regions in which a proportion of carbon is oxidized only to carbon monoxide" (WHO 2006). Breathing unsafe levels of CO lowers the amount of oxygen reaching the bloodstream and organs. This can cause chest pains and other symptoms, especially in those suffering from heart disease, and lead to hospital and emergency room admissions (EPA 2020).

*Nitrogen Dioxide* ($NO_2$) belongs to a group of gases known as oxides of nitrogen, and is by far the most dangerous of these gases. The majority of ambient $NO_2$ is formed atmospherically, when nitrogen given off in the burning of fossil fuels combines with oxygen in a reaction catalyzed by the presence of sunlight (WHO 2006). “Short-term exposures to $NO_2$ can aggravate respiratory diseases, particularly asthma, leading to respiratory symptoms, hospital admissions and emergency department visits. Long-term exposures to $NO_2$ may contribute to asthma development and potentially increase susceptibility to respiratory infections” (EPA 2020).

*Ozone* ($O_3$) is a compound occurring naturally in the stratosphere, where it protects living things from the sun’s harmful radiation. But ozone can also form as a result of “chemical reactions between oxides of nitrogen ($NO_2$) and volatile organic compounds (VOC). This happens when pollutants emitted by cars, power plants, industrial boilers, refineries, chemical plants, and other sources chemically react in the presence of sunlight” (EPA 2020). Ozone damages lung function and leads to respiratory symptoms like coughing and shortness of breath. It can also worsen asthma and lung diseases leading to increased medication use and hospital visits, as well as raise the risk of premature death from respiratory disease (EPA 2020).

*Particulate Matter* ($PM^x$) is a general term for various tiny particles of pollution from various sources that are found in the air supply. This study looks at both larger particles ($PM^{10}$), and the smaller and far more dangerous $PM^{2.5}$, which can penetrate lungs and enter the bloodstream (WHO 2006). “Exposures to PM...can cause harmful effects on the cardiovascular system including heart attacks and strokes.

These effects can result in emergency department visits, hospitalizations and, in some cases, premature death. PM exposures are also linked to harmful respiratory effects, including asthma attacks" (EPA 2020).

*Sulfur Dioxide* ($SO_2$) is produced mainly in the process of burning fuel sources that contain sulfur. Sulfur Dioxide pollution is less present in developed countries, where motor vehicle fuels undergo a refining process and industrial emissions are filtered before being released into the atmosphere. However, the burning of coal as well as unrefined gasoline leads to a higher presence of $SO_2$ in underdeveloped nations (WHO 2006). According to the EPA, "Short-term exposures to $SO_2$ are linked with respiratory effects including difficulty breathing and increased asthma symptoms… [and] have also been connected to increased emergency department visits and hospital admissions for respiratory illnesses" (2020).

## 2.2 Literature Review

A limited but growing body of literature addresses the link between air pollution and school absences. Currie et al (2009) argue in their pivotal study "Does Pollution Increase School Absences," for the use of absences as a proxy for health:

> [Absence data are] 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 offer a window on these illnesses. Moreover, there is a long tradition of using absence from school to define disability among children (Currie et al, 2009).

There are, however, limiting factors in attempting to associate air pollution with school absences. This is noted in a relatively early exploration of this relationship by Gilliland et al (2001), “The Effects of Ambient Air Pollution on School Absenteeism Due to Respiratory Illnesses”:

> Population-based studies show that absence rates vary by school, age, grade, and gender, and are affected by family structure, function, and other social factors. Although the non-health-related influences on absenteeism limit its usefulness as a measure of the adverse effects of air pollution, the majority of school absences are illness related and attributable to either respiratory infections or gastroenteritis, suggesting that illness is the dominant factor for school absenteeism (Gilliland et al., 2001).

The growing number of studies in the U.S. since 2001 which have found at least a meaningful association, if not a direct causal relationship, between pollution levels and absenteeism seem to suggest an overall trend, both in terms of specific areas of the country and over the period of time this study aims to address, namely the past 15 years.

The earliest significant study of this relationship in the United States might be Ransom and Pope (1992), a study focusing on the effects of particulate matter pollution in the Provo, Utah area. The study’s data set (1985-1990) includes a time period in which a local steel mill shuts down, offering a natural opportunity to examine the effects of a sudden drop in $PM^{10}$ pollution. (This data set is from a time before $PM^{2.5}$ was considered the standard for particulate matter measurement.)

The authors find that absences rose in the range of 54-77% when $PM^{10}$ rose from 50 to over 100 micrograms per cubic meter ($\mu g/m^3$) (Ransom & Pope, 1992).

The authors revisit this data set in their 2013 study, this time examining both particulate matter and carbon monoxide levels during the 13-month period of the steel mill shutdown. The reasoning for returning to the data is that, because the mill was a major source of particulate matter but not of carbon monoxide, it is possible to disentangle $PM^{10}$ and CO levels during the shutdown period. The results of this “natural experiment” indicate that $PM^{10}$ had a strong impact on school absences, but that CO did not, a result the authors acknowledge but do not attempt to explain (Ransom & Pope, 2013).

Two other early U.S. studies of particular note because of their focus on ozone are Chen (2000) and Gilliland et. al (2001). Like Ransom & Pope, both of these studies focus on communities in the Western U.S. -- Reno, Nevada and Southern California, respectively -- and also consider $PM^{10}$ exposure levels. (Again, $PM^{2.5}$ did not become the standard until 2006.) Both studies find that ozone levels have a significant positive effect on absenteeism. However, both studies also find that an increased presence of PM has a *negative* effect on school absences (Chen, 2000; Gilliland et al., 2001).

Later studies have posited that this unexpected might be the result of non-health factors such as socioeconomic background that correlate with pollution levels (Currie et al., 2009), or that $PM^{10}$ is difficult to distinguish from other airborne pollutants (Ransom & Pope 2013).

The most notable study in this area may be Currie et al. (2009), which focused on the effects of $PM^{10}$, ozone and CO on absenteeism in Texas schools from 1996-2001. Using a differences-in-differences-in-differences strategy in an effort to disentangle the correlating pollutants, the authors make a convincing case that CO has a direct positive effect on absenteeism. However, the effect of $PM^{10}$ and ozone is at best ambiguous (Currie et al., 2009).

One final notable study examining a subject in the U.S. is Hales et. al (2016), which considers the effect of the more sensitive $PM^{2.5}$ pollution on elementary school absenteeism in four distinct school districts in Utah. Three school districts located in areas of high pollution are evaluated for pollution level fluctuation while the fourth, which experiences far lower levels of pollution, acts as a control. As a result, the authors find a significant positive association between $PM^{2.5}$ and absences. The authors note, however, that this effect is “difficult to disentangle from other factors” (Hale et al., 2016).

# III. Data & Methodology

## 3.1 Data Information

This study considers data collected by the U.S. Census via the American Community Survey (ACS). Administered on a monthly basis, the ACS reaches over 3.5 million American households each year. The survey gathers a variety of social, demographic and economic data around subjects such as citizenship, income, educational attainment, employment, and housing characteristics.

This study uses the ACS 1-year estimate, which considers respondents only from communities (counties or equivalents) with populations of over 65,000. This study considers data from 2005 through 2018.

The variable of interest in this study is school attendance. Responses for individuals between the ages of six and 17 only are considered in order to narrow responses to those who would be expected to attend primary, middle, or high school. The survey asks whether respondents have attended school in the past three months. The counties considered in this study -- that is, those that match with the EPA collection data -- does not follow a normal distribution curve, but rather skews leftward, indicating a mode average attendance of .97, or nearly perfect attendance (see Figure 2).

Other explanatory variables besides age considered here include sex, race/ethnicity, family income, poverty ratio, and birthplace. The race/ethnicity identities considered here are: Hispanic, non-Hispanic white, Black, asian, and non-Hispanic (other). Poverty ratio for the ACS is expressed as a percentage of the national poverty threshold for a given year, depending on the size, age and composition of a given household. (Households earning more than five times the poverty rate are all rated at 501% of the poverty threshold.) Birthplace is expressed in a binary manner, with respondents indicating whether they were born in the United States or abroad.

Data on pollution levels is derived from the EPA's Air Quality Statistics Report (AQSR). The AQSR reports ambient pollution levels related to national air quality standards and includes readings for all six criteria pollutants. The report returns air pollution statistics from all available counties, for all six criteria pollutants, for one year.

The values shown are the highest reported during the year by all monitoring sites in the county or equivalent. Table 1 details measurement characteristics for the five criteria pollutants utilized in this study.

## 3.2 Methodology

As mentioned above, AQSR returns yearly measurements for all U.S. counties. However, the Census's ACS only includes responses from counties with a population of over 65,000. This results in a panel data set that is unbalanced, meaning that data does not exist for all counties in all time periods. In spite of this, the data set is still considered usable because results are returned from a representative set of U.S. counties (Figure 1).

To mitigate the effects of the unbalanced panel data set, this study considers three types of regression models: pooled OLS, random effects, and fixed effects. A pooled OLS model uses both the between group and within group variation to establish parameters. This makes it possible to pool the data and conduct an ordinary least squares dummy variable (LSDV) regression to determine whether the OLS results are consistent. In a random effects model, the individual effects are distributed randomly across the cross-sectional units and in order to capture the individual effects, and the regression model is specified with an intercept term representing an overall constant term. Whereas in a fixed effects model, the group means are fixed (non-random). Each individual has a different intercept term and the same slope parameters.

After running these three regression models it is necessary to determine which best fits the data. As a rule, a fixed effect model would be preferred due to its consistency, but a random effects model could turn out to be preferable if its consistency is found to be constant, as it is more efficient. To determine the answer, this study performs a Hausman test on the random effects model and finds the hypothesis to be rejected, therefore confirming the fixed effect model to be the preferred method model accounting for estimating the heterogeneous effects of air pollution on student school attendance.

The fixed effect regression model is

$$Y_{it} = \beta_1 X_{1,it} + \cdots + \beta_k X_{k,it} + \alpha_i + u_{it}$$

With $i$ = 1, . . . ,n and $t$ = 1,. . .,T. The $\alpha_i$ are entity-specific intercepts that capture heterogeneity across counties.

# IV. Empirical Results

## 4.1 Main Results

Four of the 13 pollution variables considered for this study were found to have some kind of significant effect on absenteeism: $NO_2$ 98th percentile; ozone 1-hr 2nd max; $SO_2$ 99th percentile; and $SO_2$ 24-hr 2nd max. (See Table 1 for an explanation of these variables.) Of the four, only ozone was found to have a positive, significant and consistent effect on absenteeism. For every unit increase in parts per million of ozone, absenteeism increased by about 3.5 per cent (Table 3).

$NO_2$ did not present initially as significant, and was only found to have a positive and significant effect when including the lag. ($NO_2$ is a constituent element of ozone, which may account for its registering a positive effect in the lag.) Both variables of $SO_2$ returned significant *negative* effects on absenteeism which disappeared when adding in the lag, and for that reason can be easily discounted. Therefore, it was determined that the ozone measurement was the lone variable warranting consideration.

Table 2 outlines our summary statistics. For our dependent variable, school attendance, we find more variation in our standard deviation within individuals over time than between individual respondents, which could indicate that external conditions are acting to suppress attendance in some years but not others. For our independent variable, ozone levels, it is the opposite: there is more variation between reporting locations than within the same reporting location over time. This is logical as different areas of the country experience wildly different levels of pollution.

Table 4 and Table 5 show the results of the OLS and LSDV regressors, respectively. The results of the LSDV match those of the fixed effects regression (Table 6), indicating that the OLS is not consistent and that fixed effect is, in fact, or preferred method.

Table 6 outlines the results of the fixed effects regression. For every unit increase in parts per million of the one hour $2^{nd}$ maximum value of ozone, absenteeism increases by about 4.2 percentage points, and is significant to a threshold of 95 percent. This result is identical when respondents identifying as Black are used as a base variable as well as when those identifying as White are used.

The result also holds up when the trend is included, which is not surprising because the trend itself is shown to be flat in all of the models run (see Figure 3). Including the lagged pollution levels also did not affect the results significantly, either alone or when included with the trend.

### 4.2 Secondary Results

As shown in Table 6, the following explanatory variables were submitted to the fixed effects regression: U.S. citizenship, place of birth, poverty ratio, total income, sex, age, and four non-White race or ethnic groupings (Black, Hispanic, Asian and other). Of these, only the poverty ratio affected a significant positive increase on the levels of absenteeism when elevated levels of ozone were present. Another finding of note: belonging to a given race or ethnic group did not seem to increase the positive effect of ozone on absenteeism. However, when including an interaction between respondents identifying as Black and the ozone data, the overall percentage of absenteeism rose to nearly 4.8 per cent.

## V. Conclusion

This paper sought to examine the relationship between ambient air pollution and school absenteeism in counties across the United States. By comparing EPA air quality readings with Census survey data, we found a possible link between heightened levels of ozone and school absences.

These results are in keeping with a number of the regionally-focused studies found in the literature. Also, absenteeism seemed to be exacerbated when family income relative to the poverty line is considered, as well as when cross-referenced with percent individuals identifying as Black.

As covered in the literature review, all papers on this subject appear to study this issue at a local or, at best, state level. More study is needed to understand this issue at the national level. This is true of individual pollutants, the presence of which are difficult to disentangle from one another, but it is especially true in terms of the veracity of data from survey respondents. For example, examining year-over-year absentee rates paints a very broad picture; more might be discovered by following a cohort of students for several years and soliciting absenteeism information on a monthly or even weekly basis. Researchers might also seek to compile absence reporting from schools directly, as is done in many of the studies in the literature. However, we see there a wide variety of reporting protocols across different school systems that would be very difficult to reconcile at the national level.

The opportunity to study subjects according to key identifiers would also be lost unless a school system was able to provide such information in a manner consistent with privacy laws. Overall, the interaction between pollution and absences when applied to similar groups across the country seems under-examined.

# VI. Bibliography

Bates, D. V. (1995). The Effects of Air Pollution on Children. Environmental Health Perspectives, 103, 49. doi:10.2307/3432345

Chen, S., Guo, C., and; Huang, X. (2018). Air Pollution, Student Health, and School Absences: Evidence from China. Journal of Environmental Economics and Management, 92, 465-497. doi:10.1016/j.jeem.2018.10.002

Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M. and Rivkin, S. G. (2009). Does Pollution Increase School Absences? Review of Economics and Statistics, 91(4), 682-694. doi:10.1162/rest.91.4.682

Environmental Protection Agency. (2020). Our Nation's Air. Retrieved July 29, 2020, from [https://gispub.epa.gov/air/trendsreport/2020/](https://gispub.epa.gov/air/trendsreport/2020/)

Gilliland, F. D., Berhane, K., Rappaport, E. B., Thomas, D. C., Avol, E., Gauderman, W. J., London, S. J., Margolis, H. G., McConnell, R., Islam, K. T., & Peters, J. M. (2001). The effects of ambient air pollution on school absenteeism due to respiratory illnesses. Epidemiology (Cambridge, Mass.), 12(1), 43–54. https://doi.org/10.1097/00001648-200101000-00009

Gottfried, M. A. (2014). Chronic Absenteeism and Its Effects on Students' Academic and Socioemotional Outcomes. Journal of Education for Students Placed at Risk (JESPAR), 19(2), 53-75. doi:10.1080/10824669.2014.962696

Grossman, Michael, and Robert Kaestner. (1997). The Effects of Education on Health. In Jere Behrman and Nancy Stacy (Eds.), The Social Benefits of Education (pp. 69–125). University of Michigan Press.

Hales, N. M., Barton, C. C., Ransom, M. R., Allen, R. T., and; Pope, C. A. (2016). A Quasi-Experimental Analysis of Elementary School Absences and Fine Particulate Air Pollution. Medicine, 95(9). doi:10.1097/md.0000000000002916

Ransom, M. R. and Pope, C. A. (1992). Elementary school absences and PM10 pollution in Utah Valley. Environmental Research, 58(1-2), 204-219. doi:10.1016/s0013-9351(05)80216-6

Ransom, M. and C.A. Pope. (2013). Air Pollution and School Absenteeism: Results from a Natural Experiment. [Working Paper]. Brigham Young University.

U.S. Census Bureau. (2020, July 21). American Community Survey (ACS). Retrieved July 29, 2020, from http://www.census.gov/programs-surveys/acs/

World Health Organization. Air quality guidelines: Global update 2005: Particulate matter, ozone, nitrogen dioxide and sulfur dioxide. (2006). Copenhagen: WHO Regional Office for Europe. Retrieved July 29, 2020, from https://www.euro.who.int/__data/assets/pdf_file/0005/78638/E90038.pdf

World Health Organization. Ambient (outdoor) air pollution. (2018, 2 May). Retrieved July 29, 2020, from https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

World Health Organization. Regional Office for Europe & European Centre for Environment and Health. (2005). Effects of air pollution on children's health and development : a review of the evidence. Copenhagen : WHO Regional Office for Europe. https://apps.who.int/iris/handle/10665/107652

# VIII. Tables

## Table 1. Pollution Variables

| Name | Description |
|---|---|
| **CO 1-hr 2nd Max** | For Carbon Monoxide, the 2nd highest 1-hour measurement in the year. |
| **CO 8-hr 2nd Max** | For Carbon Monoxide, the 2nd highest non-overlapping 8-hour average in the year. |
| **NO2 98th %ile** | For Nitrogen Dioxide, the 98th percentile of the daily max 1-hour measurements/year. |
| **NO2 Annual Mean** | For Nitrogen Dioxide, the annual mean of all the 1-hour measurements in the year. |
| **O3 1-hr 2nd Max** | For Ozone, the 2nd highest daily max 1-hour measurement in the year. |
| **O3 8-hr 4th Max** | For Ozone, the 4th highest daily max 8-hour average in the year. |
| **SO2 99th %ile** | For Sulfur Dioxide, the 99th percentile of the daily max 1-hour measurements in the year. |
| **SO2 24-hr 2nd Max** | For Sulfur Dioxide, the 2nd highest 24-hour average measurement in the year. |
| **SO2 Annual Mean** | For Sulfur Dioxide, the annual mean of all the 1-hour measurements in the year. |
| **PM2.5 98th %ile** | For PM2.5, the 98th percentile of the daily average measurements in the year. |
| **PM2.5 Wtd Mean** | For PM2.5, the Weighted Annual Mean (mean weighted by calendar quarter) for the year. |
| **PM10 24-hr 2nd Max** | For PM10, the 2nd highest 24-hour average measurement in the year. |
| **PM10 Annual Mean** | For PM10, the Weighted Annual Mean (mean weighted by calendar quarter) for the year. |

(EPA, 2020)

# Table 2. Summary Statistics

| Variable | | Mean | Std. Dev. | Min | Max | Observations |
|---|---|---|---|---|---|---|
| | overall | 207.5977 | 124.4483 | 1 | 423 | N = 4830 |
| Group County Code | between | - | 122.2538 | 1 | 423 | n = 423 |
| | within | - | 0 | 207.5977 | 207.5977 | T-bar = 11.4184 |
| | overall | 2011.719 | 4.022616 | 2005 | 2018 | N = 4830 |
| Year | between | - | 2.015191 | 2005.5 | 2017.5 | n = 423 |
| | within | - | 3.7279 | 2002.969 | 2019.147 | T-bar = 11.4184 |
| | overall | .9773389 | .0119869 | .8076923 | 1 | N = 4830 |
| School Attendance | between | - | .0063957 | .9412652 | .9913538 | n = 423 |
| | within | - | .0102916 | .830902 | 1.01684 | T-bar = 11.4184 |
| | overall | .0889972 | .0159773 | .04 | .18 | N = 4308 |
| Ozone 2nd Max 1-hr | between | - | .0123326 | .05 | .15 | n = 375 |
| | within | - | .0099364 | .0611401 | .1918544 | T-bar = 11.488 |
| | overall | .4737258 | .2447573 | 0 | 1 | N = 4715 |
| US Citizen | between | - | .1601206 | .0833333 | 1 | n = 423 |
| | within | - | .195782 | -.3403768 | 1.236764 | T-bar = 11.1466 |
| | overall | .9519085 | .0325904 | .7938597 | 1 | N = 4830 |
| US Born | between | - | .0284032 | .8412163 | .9943255 | n = 423 |
| | within | - | .0139197 | .8577362 | 1.021587 | T-bar = 11.4184 |
| | overall | 290.5595 | 52.62935 | 145.4342 | 440.1619 | N = 4830 |
| Poverty Ratio | between | - | 48.88919 | 163.5933 | 425.6913 | n = 423 |
| | within | - | 15.15705 | 210.5143 | 353.4763 | T-bar = 11.4184 |
| | overall | 72138.76 | 14409.26 | 33471.7 | 124769.3 | N = 4830 |
| Total Family Income | between | - | 13211.7 | 40623.79 | 116115.9 | n = 423 |
| | within | - | 5420.066 | 51948.05 | 101450 | T-bar = 11.4184 |
| | overall | .4870626 | .0267554 | .3566433 | .6081081 | N = 4830 |
| Female | between | - | .0094604 | .4462617 | .5222022 | n = 423 |
| | within | - | .0254345 | .3631359 | .60381 | T-bar = 11.4184 |
| | overall | 11.64105 | .2345155 | 10.45133 | 12.64368 | N = 4830 |
| Age | between | - | .1328161 | 10.97422 | 12.26054 | n = 423 |
| | within | - | .1980112 | 10.74137 | 12.51225 | T-bar = 11.4184 |
| | overall | .8329729 | .1776989 | .0062893 | 1 | N = 4830 |
| Hispanic | between | - | .1684334 | .0295122 | .993063 | n = 423 |
| | within | - | .0274296 | .6981585 | .9563043 | T-bar = 11.4184 |
| | overall | .0044243 | .0069391 | 0 | .0805987 | N = 4830 |
| Black | between | - | .0052294 | 0 | .0533374 | n = 423 |
| | within | - | .0043391 | -.0156431 | .0432476 | T-bar = 11.4184 |
| | overall | .0028854 | .0052951 | 0 | .0789474 | N = 4830 |
| Asian | between | - | .0036552 | 0 | .0309535 | n = 423 |
| | within | - | .0037913 | -.0232372 | .0508793 | T-bar = 11.4184 |
| | overall | .0538998 | .0603389 | 0 | .4548998 | N = 4830 |
| Other | between | - | .0537415 | 0 | .3992056 | n = 423 |
| | within | - | .0216991 | -.0864119 | .2843074 | T-bar = 11.4184 |
| | overall | .1058176 | .1315663 | 0 | .9449378 | N = 4830 |
| White | between | - | .1228069 | .0046494 | .8707292 | n = 423 |
| | within | - | .0303832 | -.1112009 | .2672144 | T-bar = 11.4184 |

# Table 3. Estimated Effects of $NO_2$, $O_3$ and $SO_2$ on Average School Attendance

| VARIABLES | (1) School Attendance | (2) School Attendance | (3) School Attendance | (4) School Attendance |
|---|---|---|---|---|
| NO2 98th Percentile | -7.93e-06 | | | |
| | (3.25e-05) | | | |
| Lag NO2 98th Percentile | -8.89e-05** | | | |
| | (4.41e-05) | | | |
| O3 1-hr 2nd Max | | -0.0354* | | |
| | | (0.0212) | | |
| Lag O3 1-hr 2nd Max | | -0.0133 | | |
| | | (0.0185) | | |
| SO2 99th Percentile | | | 3.49e-06*** | |
| | | | (5.33e-07) | |
| Lag SO2 99th Percentile | | | -5.81e-07 | |
| | | | (7.33e-07) | |
| SO2 24-hr 2nd Max | | | | 1.47e-05** |
| | | | | (6.13e-06) |
| Lag SO2 24-hr 2nd Max | | | | -2.01e-06 |
| | | | | (5.47e-06) |
| US Citizen | -0.000692 | -0.000715 | -0.00217 | -0.00218 |
| | (0.00267) | (0.00133) | (0.00212) | (0.00212) |
| US Born | 0.0360* | 0.0126 | 0.00825 | 0.00835 |
| | (0.0199) | (0.0131) | (0.0285) | (0.0284) |
| Poverty Ratio | -4.41e-05** | -4.02e-05*** | -3.27e-05 | -3.25e-05 |
| | (2.16e-05) | (1.34e-05) | (2.51e-05) | (2.50e-05) |
| Female | -0.0240* | 0.0133 | 0.00895 | 0.00900 |
| | (0.0126) | (0.00819) | (0.0134) | (0.0134) |
| Age | -0.00107 | -0.00233** | -0.00327 | -0.00323 |
| | (0.00223) | (0.00115) | (0.00220) | (0.00219) |
| Hispanic | 0.0155 | -0.000957 | 0.0156 | 0.0154 |
| | (0.0135) | (0.00955) | (0.0140) | (0.0140) |
| Black | 0.0230 | -0.0298 | 0.0646 | 0.0648 |
| | (0.0603) | (0.0443) | (0.0585) | (0.0586) |
| Asian | 0.140** | 0.110** | -0.0126 | -0.0218 |
| | (0.0578) | (0.0495) | (0.0724) | (0.0781) |
| Other | 0.00583 | -0.00768 | -0.00896 | -0.00920 |
| | (0.00872) | (0.00720) | (0.0130) | (0.0130) |
| Linear Trend | -1.21e-05 | -4.34e-05 | 0.000151 | 0.000154 |
| | (0.000126) | (9.20e-05) | (0.000103) | (0.000103) |
| Observations | 1,354 | 3,818 | 1,462 | 1,464 |
| No. of Group County Code | 169 | 372 | 192 | 192 |
| Fixed Effect | YES | YES | YES | YES |
| Lag | YES | YES | YES | YES |
| Year Dummy | NO | NO | NO | NO |
| Linear Trend | YES | YES | YES | YES |

Some controls were omitted from the table

Robust standard errors in parentheses - Clustered s.e.

*** p<0.01, ** p<0.05, * p<0.1

# Table 4. Estimated Pooled OLS Effects of $O_3$ 1-hr 2nd Max on Average School Attendance

| VARIABLES | (1) School Attendance | (2) School Attendance | (3) School Attendance | (4) School Attendance | (5) School Attendance | (6) School Attendance |
|---|---|---|---|---|---|---|
| O3 1-hr 2nd Max | -0.00955 | -0.0110 | -0.0101 | -0.00870 | -0.0120 | -0.0115 |
| | (0.0146) | (0.0177) | (0.0189) | (0.0154) | (0.0180) | (0.0176) |
| O3 1-hr 2nd Max x Black | | 0.355 | 0.344 | | | |
| | | (2.084) | (2.095) | | | |
| Linear Trend | | | 1.88e-05 | 1.91e-05 | -3.51e-05 | |
| | | | (0.000101) | (0.000101) | (0.000102) | |
| Lag O3 1-hr 2nd Max | | | | | 0.00615 | 0.00765 |
| | | | | | (0.0140) | (0.0141) |
| US Citizen | -0.000802 | -0.000801 | -0.000816 | -0.000817 | -0.000721 | -0.000750 |
| | (0.00111) | (0.00111) | (0.00110) | (0.00110) | (0.00115) | (0.00115) |
| US Born | 0.00456 | 0.00460 | 0.00465 | 0.00460 | 0.00702 | 0.00710 |
| | (0.0111) | (0.0111) | (0.0111) | (0.0110) | (0.0112) | (0.0112) |
| Poverty Ratio | 2.18e-05 | 2.19e-05 | 2.47e-05 | 2.47e-05 | 4.21e-05* | 4.70e-05*** |
| | (1.46e-05) | (1.46e-05) | (2.25e-05) | (2.25e-05) | (2.43e-05) | (1.68e-05) |
| Total Family Income | 1.28e-07*** | 1.28e-07** | 1.17e-07 | 1.17e-07 | 5.83e-08 | 3.90e-08 |
| | (4.95e-08) | (4.96e-08) | (8.21e-08) | (8.22e-08) | (8.81e-08) | (5.77e-08) |
| Female | 0.00309 | 0.00309 | 0.00307 | 0.00306 | 0.00822 | 0.00818 |
| | (0.00754) | (0.00754) | (0.00754) | (0.00754) | (0.00816) | (0.00816) |
| Age | -0.000481 | -0.000485 | -0.000480 | -0.000475 | -0.000790 | -0.000780 |
| | (0.00121) | (0.00120) | (0.00120) | (0.00120) | (0.00130) | (0.00131) |
| Hispanic | -0.00948*** | -0.00951*** | -0.00951*** | -0.00948*** | -0.00941*** | -0.00941*** |
| | (0.00315) | (0.00316) | (0.00315) | (0.00315) | (0.00319) | (0.00318) |
| Black | -0.0294 | -0.0606 | -0.0605 | -0.0304 | -0.0335 | -0.0352 |
| | (0.0353) | (0.188) | (0.188) | (0.0370) | (0.0353) | (0.0339) |
| Asian | 0.0172 | 0.0172 | 0.0174 | 0.0174 | 0.0174 | 0.0180 |
| | (0.0547) | (0.0547) | (0.0545) | (0.0544) | (0.0582) | (0.0582) |
| Other | 0.00389 | 0.00374 | 0.00382 | 0.00397 | 0.00603 | 0.00609 |
| | (0.00698) | (0.00699) | (0.00696) | (0.00695) | (0.00693) | (0.00693) |
| Observations | 4,221 | 4,221 | 4,221 | 4,221 | 3,818 | 3,818 |
| R-squared | 0.059 | 0.059 | 0.059 | 0.059 | 0.059 | 0.059 |
| Fixed Effect | NO | NO | NO | NO | NO | NO |
| Linear Trend | NO | NO | YES | YES | YES | NO |
| Lag | NO | NO | NO | NO | YES | YES |
| Interaction | NO | YES | YES | NO | NO | NO |

Robust standard errors in parentheses - Clustered s.e.
*** p<0.01, ** p<0.05, * p<0.1

# Table 5. Estimated OLS - (LSDV) Effects of $O_3$ 1-hr 2nd Max on Average School Attendance

| VARIABLES | (1) School Attendance | (2) School Attendance | (3) School Attendance | (4) School Attendance | (5) School Attendance | (6) School Attendance |
|---|---|---|---|---|---|---|
| O3 1-hr 2nd Max | -0.0422** | -0.0476* | -0.0487* | -0.0434* | -0.0356 | -0.0355* |
| | (0.0213) | (0.0247) | (0.0258) | (0.0222) | (0.0224) | (0.0214) |
| O3 1-hr 2nd Max x Black | | 1.322 | 1.316 | | | |
| | | (2.380) | (2.375) | | | |
| Linear Trend | | | -2.52e-05 | -2.58e-05 | -2.56e-06 | |
| | | | (0.000119) | (0.000119) | (0.000130) | |
| Lag O3 1-hr 2nd Max | | | | | -0.0130 | -0.0128 |
| | | | | | (0.0194) | (0.0182) |
| US Citizen | -0.000897 | -0.000894 | -0.000871 | -0.000873 | -0.000715 | -0.000717 |
| | (0.00135) | (0.00135) | (0.00135) | (0.00134) | (0.00140) | (0.00141) |
| US Born | 0.0170 | 0.0169 | 0.0173 | 0.0173 | 0.0123 | 0.0122 |
| | (0.0129) | (0.0130) | (0.0131) | (0.0131) | (0.0138) | (0.0137) |
| Poverty Ratio | -6.22e-05*** | -6.17e-05*** | -6.50e-05*** | -6.56e-05*** | -3.04e-05 | -3.00e-05 |
| | (1.97e-05) | (1.98e-05) | (2.42e-05) | (2.41e-05) | (2.60e-05) | (2.22e-05) |
| Total Family Income | 5.71e-08 | 5.53e-08 | 7.12e-08 | 7.32e-08 | -4.23e-08 | -4.38e-08 |
| | (6.53e-08) | (6.58e-08) | (9.56e-08) | (9.55e-08) | (1.02e-07) | (7.62e-08) |
| Female | 0.00799 | 0.00800 | 0.00803 | 0.00803 | 0.0133 | 0.0133 |
| | (0.00790) | (0.00790) | (0.00789) | (0.00790) | (0.00863) | (0.00863) |
| Age | -0.00204* | -0.00206* | -0.00206* | -0.00205* | -0.00233* | -0.00233* |
| | (0.00111) | (0.00111) | (0.00111) | (0.00111) | (0.00121) | (0.00121) |
| Hispanic | -0.00335 | -0.00347 | -0.00408 | -0.00398 | -0.000832 | -0.000776 |
| | (0.00899) | (0.00902) | (0.00952) | (0.00950) | (0.0100) | (0.00951) |
| Black | -0.0315 | -0.145 | -0.145 | -0.0316 | -0.0291 | -0.0291 |
| | (0.0440) | (0.218) | (0.217) | (0.0439) | (0.0464) | (0.0464) |
| Asian | 0.0952* | 0.0948* | 0.0946* | 0.0950* | 0.111** | 0.111** |
| | (0.0494) | (0.0494) | (0.0494) | (0.0494) | (0.0523) | (0.0524) |
| Other | -0.00575 | -0.00573 | -0.00589 | -0.00591 | -0.00779 | -0.00778 |
| | (0.00723) | (0.00725) | (0.00724) | (0.00722) | (0.00761) | (0.00759) |
| Observations | 4,221 | 4,221 | 4,221 | 4,221 | 3,818 | 3,818 |
| R-squared | 0.280 | 0.280 | 0.280 | 0.280 | 0.290 | 0.290 |
| Fixed Effect | NO | NO | NO | NO | NO | NO |
| Linear Trend | NO | NO | YES | YES | YES | NO |
| Lag | NO | NO | NO | NO | YES | YES |
| Interaction | NO | YES | YES | NO | NO | NO |

No.of 423 Countycodes were included in the regression but omitted from the table.
Robust standard errors in parentheses - Clustered s.e.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

# Table 6. Estimated Fixed Effects of $O_3$ 1-hr 2$^{nd}$ Max on Average School Attendance

| VARIABLES | (1) School Attendance | (2) School Attendance | (3) School Attendance | (4) School Attendance | (5) School Attendance | (6) School Attendance | (7) School Attendance |
|---|---|---|---|---|---|---|---|
| O3 1-hr 2$^{nd}$ Max | -0.0422** | -0.0422** | -0.0476** | -0.0487** | -0.0434** | -0.0356* | -0.0355* |
| | (0.0203) | (0.0203) | (0.0236) | (0.0246) | (0.0212) | (0.0213) | (0.0203) |
| O3 1-hr 2$^{nd}$ Max x Black | | | 1.322 | 1.316 | | | |
| | | | (2.271) | (2.267) | | | |
| Linear Trend | | | | -2.52e-05 | -2.58e-05 | -2.56e-06 | |
| | | | | (0.000114) | (0.000114) | (0.000123) | |
| Lag O3 1-hr 2$^{nd}$ Max | | | | | | -0.0130 | -0.0128 |
| | | | | | | (0.0184) | (0.0173) |
| US Citizen | -0.000897 | -0.000897 | -0.000894 | -0.000871 | -0.000873 | -0.000715 | -0.000717 |
| | (0.00129) | (0.00129) | (0.00129) | (0.00128) | (0.00128) | (0.00133) | (0.00134) |
| US Born | 0.0170 | 0.0170 | 0.0169 | 0.0173 | 0.0173 | 0.0123 | 0.0122 |
| | (0.0124) | (0.0124) | (0.0124) | (0.0125) | (0.0125) | (0.0131) | (0.0130) |
| Poverty Ratio | -6.22e-05*** | -6.22e-05*** | -6.17e-05*** | -6.50e-05*** | -6.56e-05*** | -3.04e-05 | -3.00e-05 |
| | (1.88e-05) | (1.88e-05) | (1.89e-05) | (2.31e-05) | (2.30e-05) | (2.47e-05) | (2.11e-05) |
| Total Family Income | 5.71e-08 | 5.71e-08 | 5.53e-08 | 7.12e-08 | 7.32e-08 | -4.23e-08 | -4.38e-08 |
| | (6.23e-08) | (6.23e-08) | (6.29e-08) | (9.12e-08) | (9.11e-08) | (9.73e-08) | (7.24e-08) |
| Female | 0.00799 | 0.00799 | 0.00800 | 0.00803 | 0.00803 | 0.0133 | 0.0133 |
| | (0.00754) | (0.00754) | (0.00754) | (0.00753) | (0.00754) | (0.00820) | (0.00820) |
| Age | -0.00204* | -0.00204* | -0.00206* | -0.00206* | -0.00205* | -0.00233** | -0.00233** |
| | (0.00106) | (0.00106) | (0.00106) | (0.00106) | (0.00106) | (0.00115) | (0.00115) |
| Hispanic | 0.0281 | -0.00335 | -0.00347 | -0.00408 | -0.00398 | -0.000832 | -0.000776 |
| | (0.0404) | (0.00858) | (0.00861) | (0.00909) | (0.00906) | (0.00952) | (0.00903) |
| Asian | 0.127* | 0.0952** | 0.0948** | 0.0946** | 0.0950** | 0.111** | 0.111** |
| | (0.0648) | (0.0471) | (0.0471) | (0.0472) | (0.0472) | (0.0497) | (0.0497) |
| Other | 0.0257 | -0.00575 | -0.00573 | -0.00589 | -0.00591 | -0.00779 | -0.00778 |
| | (0.0417) | (0.00690) | (0.00692) | (0.00691) | (0.00689) | (0.00723) | (0.00721) |
| Black | | -0.0315 | -0.145 | -0.145 | -0.0316 | -0.0291 | -0.0291 |
| | | (0.0420) | (0.208) | (0.208) | (0.0419) | (0.0440) | (0.0441) |
| White | 0.0315 | | | | | | |
| | (0.0420) | | | | | | |
| Observations | 4,221 | 4,221 | 4,221 | 4,221 | 4,221 | 3,818 | 3,818 |
| R-squared | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 | 0.011 | 0.011 |
| No. of Group County Code | 375 | 375 | 375 | 375 | 375 | 372 | 372 |
| Fixed Effect | YES | YES | YES | YES | YES | YES | YES |
| Linear Trend | NO | NO | NO | YES | YES | YES | NO |
| Lag | NO | NO | NO | NO | NO | YES | YES |
| Interaction | NO | NO | YES | YES | NO | NO | NO |

Robust standard errors in parentheses – Clustered s.e.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

# IX. Figures

Figure 1. Sample Counties

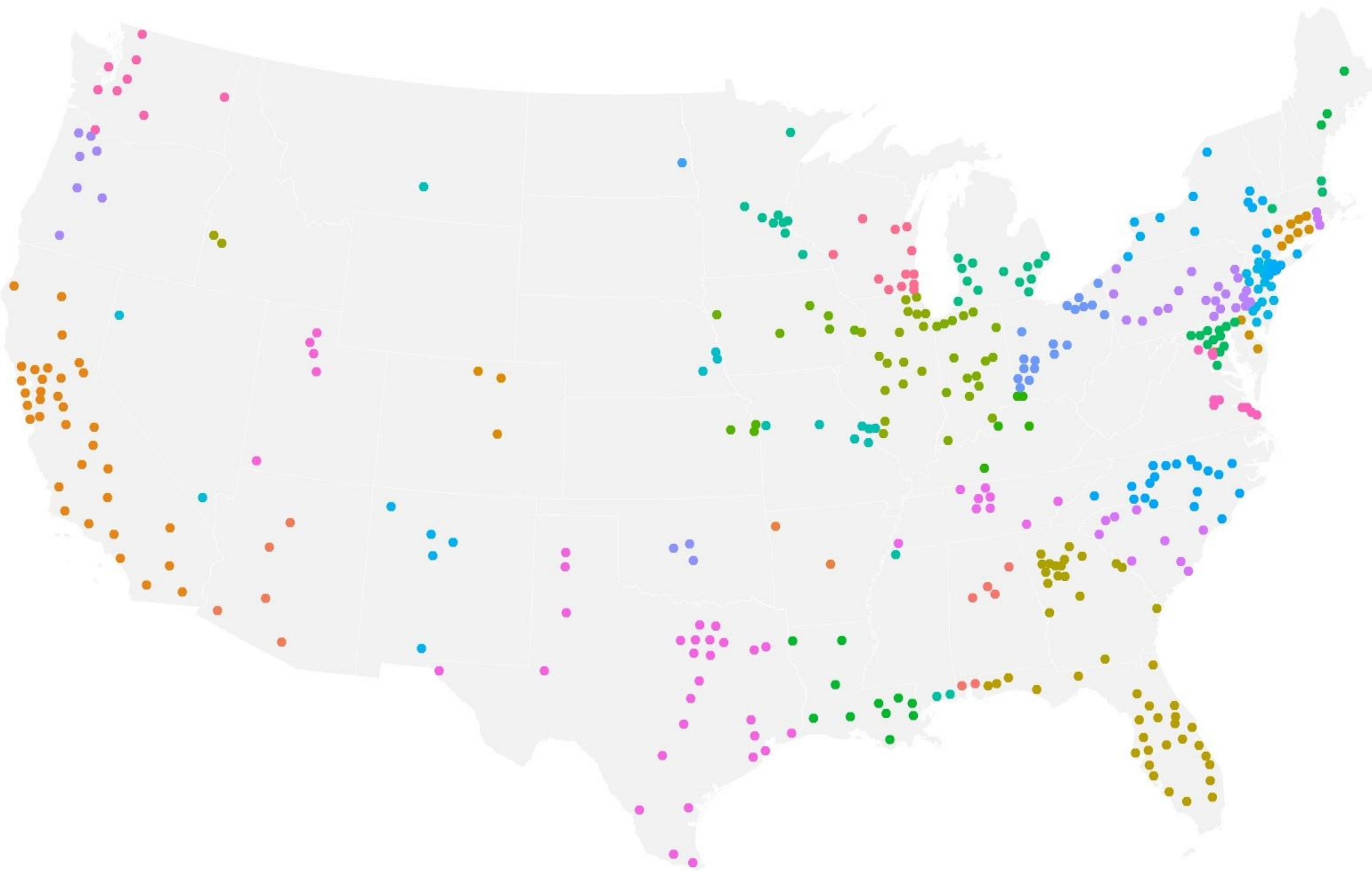
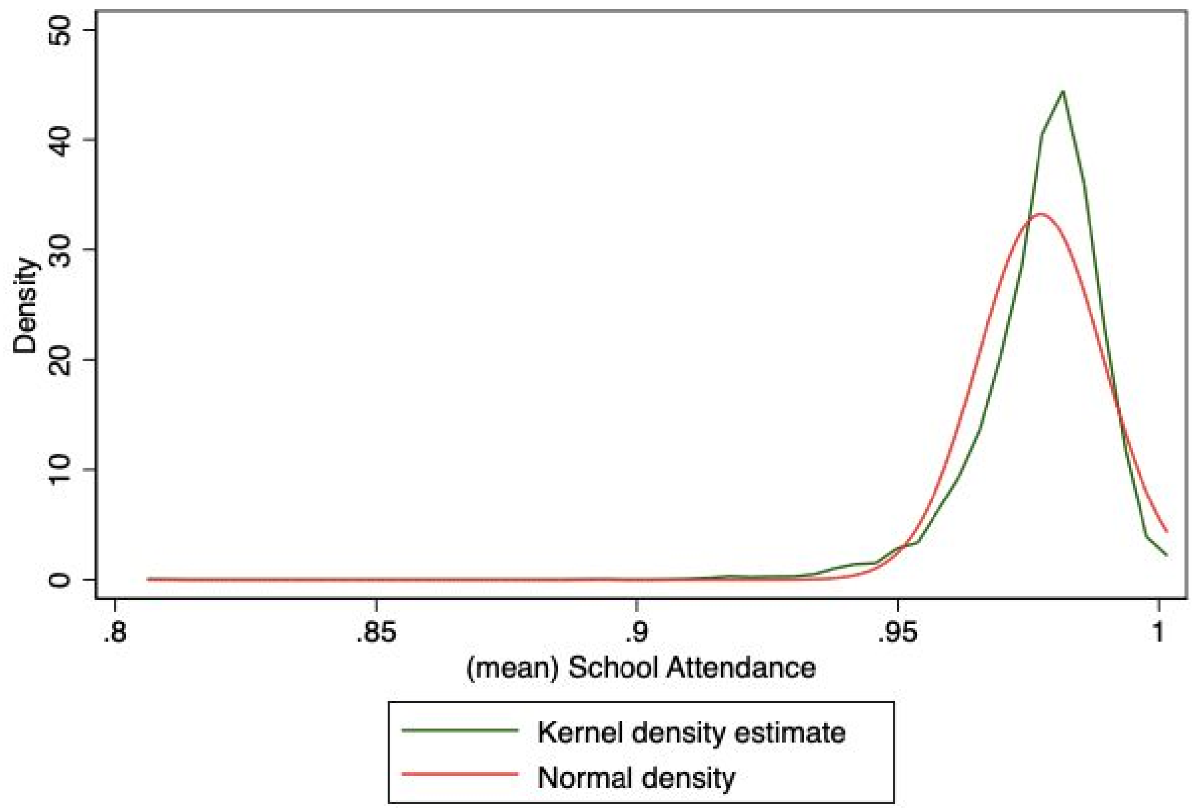


Figure 2. Sample Distribution




kernel = epanechnikov, bandwidth = 0.0016

# Figure 3. Estimated Effect of $O_3$ 1-hr 2nd Max on Average School Attendance with Linear Trend