Tell Me How I’m Supposed to Breed with No Air: Air Pollution and Fertility in the U.S.

Margaret Kallus

Honors Thesis

4/19/2019

Abstract: Exposure to air pollution during pregnancy is known to cause adverse infant health outcomes. However, less is known about the effect of pollution on fertility. In this paper, I use EPA air quality monitor data and NCHS vital statistics from 1980–2002 to examine associations between air pollution levels and race-specific county-level fertility nine months later. I use an instrumental variables framework to provide an alternate set of coefficients to OLS for fine particulate matter. With a p-value of 0.12, I find that a one standard deviation increase in fine particulate matter (PM 10) causes a roughly 7% reduction in county fertility. This is not robust to population weighting, suggesting possible treatment effect heterogeneity.

I. Introduction

There is a large literature documenting the negative effect of air pollution on a variety of health outcomes. Though average ambient levels of most air pollutants have been declining in the U.S. over the past few decades, increasing proportions of the population live in more polluted urban areas, resulting in an ambiguous change to net exposure.[1] A much smaller literature examines the effect of air pollution on fecundity and fertility. This issue is particularly relevant as the average age of first-time mothers and the percentage of children born to women older than 35 both increase. Control over fertility is clearly valued at the individual level, and difficulty conceiving affects 20 percent of American couples.[2] Fertility is also critical at the regional and societal level, since birth patterns determine future labor market and demographic patterns. Since some populations (particularly nonwhite, low SES) are more exposed to air pollution residentially, impacts on health would also have inequality implications. Understanding the effects of pollutants on fertility is critical to inform locational decisions for industry, roadways, and housing that will affect generations to come.

One of the only papers that considers the causal effect of air pollution on fertility is Clay et al. (2018), which looks at the effects of decreases in airborne lead caused by the phase-out of lead from gasoline. The authors find a 6.7 percent increase in annual mean fertility resulting from the average decrease in county-level airborne lead from 1978–1988. I add to this literature by examining the relationship between other common pollutants, specifically fine particulate matter (PM 10), on fertility at the county level from the 1980s to the 2000s.[3] Following Chay and

[1] Statistic from Center for Disease Control and Prevention Morbidity and Mortality Weekly Report, 2017.
[2] Gnoth et al., 2005. About 20 percent of couples are unable to conceive within 6 menstrual cycles, 10 percent within 12 menstrual cycles, and 5 percent within 48 menstural cycles.
[3] I look at PM 10 because this is the only regulatory change that affects a meaningful number of counties and for which I observe sufficient data before and after.

Greenstone (2005), Bishop et al. (2018), and Clay et al. (2018), I also use variation in pollution induced by the Clean Air Act in 1970 to estimate a causal effect of particulate matter on county-level fertility during this time.

II. Background

Overview of Existing Studies

Air pollution is known to cause ailments including cardiovascular disease, stroke, cancer, asthma, dementia, and premature birth (Carre et al. 2017). To the extent that these illnesses manifest in people of childbearing age, some of these conditions could have the added effect of reducing fertility at least temporarily and would be picked up in my estimates. An understanding of the validity of independent biological pathways through which air pollution reduces fertility is therefore important when assessing the relevance of this investigation. The medical literature has studied possible pathways through which pollutant exposure may cause a direct reduction in fertility and reproductive fitness. Mice and rats exposed to ambient levels of outdoor urban air pollution are shown to exhibit decreased fecundity (Mohallem et al. 2005). Additionally, examinations of women receiving in vitro fertilization treatments reveal that exposure to nitrogen dioxide, ozone, and particulate matter is associated with increased rates of miscarriage and irregular cell allocation patterns early in an embryo's development (Legro et al. 2010). Researchers have also noted associations between exposure to sulfur dioxide and particulate matter and unhealthy sperm motility and morphology, both of which impact the chance of pregnancy (Selevan et al. 2000).[4] Fewer and less rigorous studies have looked at the female

[4] Sperm motility is the sperm's ability to move efficiently, and is measured as the percent of sperm that have normal tail movement under a microscope. Sperm morphology is the shape and size of the sperm, and is measured as the percent of sperm that appear normal under a microscope. (Amelar et al., 1980).

gamete, but those that have find an association between air pollutant exposure and altered ovulation processes (Tomei et al. 2006).

This research has three primary weaknesses when interpreting findings in practice. First, people can change their behavior to mitigate the circumstances of their environment. For instance, they can reduce their exposure to pollution and thus the risk of adverse health outcomes by staying indoors or living in low-pollution areas, or they can change their sexual behavior to increase the odds of becoming pregnant. The effective impact of pollutants may therefore be different from their estimated effects in randomized control trials. Furthermore, studies on animals cannot always be extrapolated to humans. These factors limit the practical applicability of laboratory studies. Medical, non-laboratory studies of humans, though based on real-world levels of exposure, are often missing important variables such as health-seeking behavior and habits like tobacco consumption, or are not able to isolate the effect of pollution exposure (Carre et al., 2017).[5] This suggests that individual-level studies of association may not be appropriate for measuring the health impacts of air pollution. Socioeconomic status, education, race/ethnicity, level of air pollution exposure, and fertility are all correlated, preventing associative literature from isolating total causal effects of pollution on fertility.

Summary of Pollutants

The specific effects and mechanisms of different pollutants can be confusing and technical. Here, I provide a summary of the main effects and sources of each pollutant in this

[5] For instance, Thurston et al. (2000) compares women who go into careers that require pollution exposure to those who find jobs that do not require high levels of pollutant exposure. Selevan et al. (2000) examines men in two regions of the Czech Republic, one of which is industrial and one of which is rural. Systematic differences between the people in those careers/regions, which almost certainly exist, prevent a causal interpretation of the association between pollution and the desired outcome.

project and when possible contextualize the sensory impact of their presence. First, a note on units. Particulate matter is measured in micrograms per cubic meter; this is roughly equivalent to parts per billion, the unit of nitrogen dioxide and sulfur dioxide. Ozone and carbon monoxide are measured in parts per million.[6] A concentration of 3 ppm is equivalent to saying that in a million molecules of air, three would be expected to be of that pollutant. Most ozone exposure measured on EPA air quality monitors is less than 1 ppm, which seems quite small. For context, acetone (nail polish remover) can be smelled starting at 0.40 ppm, ammonia can be smelled starting at 0.043 ppm, and chlorine can be smelled starting at 0.021 ppm (American Industrial Hygiene Association, 2013).

Particulate Matter

PM 10 measures the total mass of airborne solid or liquid particles less than 10 micrometers or less in diameter, and PM 2.5 measures the total mass of these particles less than 2.5 micrometers in diameter (for context, the width of a human hair is about 50 micrometers). These particles are generated in two main ways: mechanical disruption (e.g., kicking up dust when driving over an unpaved road, construction sites, wildfires) and chemical reactions in the atmosphere. These particles enter the body mainly through the respiratory system; finer particles are able to penetrate the body further through the alveoli, and thus have more severe health consequences. To help contextualize this, Guo and Mosley (2000) measured the particulate matter released in a small laboratory when burning commercial candles. There was a lot of variety in the amount of particles released; one 9-wick paraffin candle burned in released 41 micrograms per hour per wick; another released 3120 micrograms per hour per wick. Particulate

[6] Metrics found in this table: https://www.epa.gov/criteria-air-pollutants/naaqs-table.

matter surged in both cases when the flame was put out; that little plume of smoke from blowing out a candle is an example of particulate matter (Boguski, 2006).

Ozone (O_3)

Ozone has a weaker association with fertility than particulate matter, though in a study of women undergoing IVF exposure after embryo implantation was shown to decrease the odds of live birth (Legro et al., 2010).[7] Concentrations of ozone are about twice as high outdoors as compared to indoors; fine particulate matter is only about 5 percent lower inside, and is harder to avoid via marginal actions (Laumbach et al., 2015). Ozone exists naturally in the stratosphere, and blocks some of the stronger UV rays from entering the Earth's atmosphere (the chlorofluorocarbon "ozone hole" is caused by the depletion of this stratospheric ozone). However, at the ground level, ozone can cause a variety of respiratory issues and can harm ecosystems. Ground-level ozone forms when nitrogen oxides and volatile organic compounds react in the presence of sunlight, which is why seasonal variation is particularly pronounced for this pollutant. Industrial facilities and car exhaust are some of the sources of the pollutants that become ground-level ozone.

Sulfur Dioxide (SO_2)

Sulfur dioxide is released from burning fossil fuels and copper mining; very small amounts are also released naturally in geological processes such as volcano eruptions, but 99

[7] The EPA actually releases two separate files on particulate matter: one file with observations that use EPA-approved methods, and another file that produces data that is useful enough to report but gathered with a collection method that the EPA is not allowed to use to inform decisions. For counties that have observations of each type, the median difference between the measures is 0.2 micrograms per cubic meter; this is such a small percentage of the mean of the measures that I currently construct one PM 2.5 measure from these two files. If a county-month has an observation using the approved measures, this is the estimate I use; if not, I use the non-EPA measure. The EPA-approved file begins in 1997, while the other file begins in 1988.

percent is manmade.[8] Zou et al. (2011) found that exposure to SO_2 during pregnancy increases the probability of low birthweight among infants born to mothers older than 35. Dales et al. (2004) found that exposure to this sulfur dioxide, nitrogen dioxide, and carbon monoxide increase the risk of Sudden Infant Death Syndrome in the short term. To the best of my knowledge, there are no studies linking fertility and SO_2, though Liu et al. (2017) linked SO_2 exposure and decreased sperm quality. Levels of sulfur dioxide declined a lot during this time period, but standards for sulfur dioxide were not updated between 1980 and 2002.[9] Controlling for SO_2 levels should account for the impact of this decrease.

Nitrogen Dioxide (NO_2)

Road traffic, coal-burning appliances, and tobacco smoke contribute to NO_2 production and exposure (Jarvis et al, 2010). Legro et al. (2010) showed, in a study of women undergoing IVF, that increased exposure to nitrogen dioxide during pregnancy is associated with a decreased chance of pregnancy. In another study, a 10 ppb increase in NO_2 levels was associated with a 16 percent increase in the chance of spontaneous pregnancy loss (Leiser et al., 2019). Exposure to NO_2 is thought to decrease sperm cell counts in males, though this was shown in an experiment on rats (Watanabe, 2005). Levels of nitrogen dioxide declined a lot during this time period, but standards for nitrogen dioxide were not updated between 1980 and 2002.[10] Controlling for NO_2 levels should account for the impact of this decrease.

Carbon Monoxide (CO)

---

[8] Public Health Statement, The Agency for Toxic Substances and Disease Registry (1998).
[9] Sulfur Dioxide Trends, EPA.
[10] Nitrogen Dioxide Trends, EPA.

Carbon monoxide is a colorless, odorless gas released by burning organic compounds like fossil fuels. It is also released in vehicle exhaust, tobacco smoke, and wildfires.[11] Currie (2009) finds large negative causal effects on infant health for children exposed in utero to carbon monoxide from vehicular exhaust, using individual-level data of women living near major roads as EZ-Pass was introduced in the 1990s. These effects were larger for smokers and women over 35. Exposure to high levels of carbon monoxide increases the likelihood of miscarriage, though counties do not have ambient CO levels high enough to trigger this in a meaningful way.[12] A study of women undergoing IVF found CO exposure to be associated with decreased pregnancy odds (Choe et al., 2018). Standards for carbon monoxide were not updated between 1980 and 2002.

III. Data

Birth and Population Data

For birth and population data, I use NCHS natality data and the National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER)'s county-level population estimates from 1980–2002. I drop Alaska and Hawaii from all datasets, as geographic coding schemes change inconsistently over time in these states.

NCHS natality data provides individual-level births, which I code to geographic location using the mother's county of residence to provide an estimate of in-utero exposure. This data also includes information on race, defined as the mother's race (white, black, other). I do not compute estimates for women who are neither white nor black, since this category refers to different populations in different counties. I aggregate all data to the county-race-month level.

[11] NIH ToxTown, Carbon Monoxide.
[12] The National Institute for Occupational Safety and Health, "Smoke and Byproducts of Burning."

From 1980–1988, births are available for all U.S. counties. From 1989–2002, births are only available for counties with a population greater than 100,000. Further explanation of this NCHS data can be found in the Appendix.

SEER annual population data for 1980–2002 includes population data broken counts by age, sex, and race. The number of women aged 15-44 of a given race in a county-year comprises the denominator of fertility rate, my outcome of interest.[13]

Notably, I use *annual* population data to compute *monthly* fertility rates. This is concerning if systematic seasonal migration of women 15-44 is correlated with seasonal pollution patterns, though this does not seem terribly likely. Using an annual denominator for this age group is not without precedent; this annual population estimate is also used by the BLS to compute monthly employment statistics. The fertility rate is constructed as follows:

$$Fertility\ rate_{mr} = \frac{births_{m,r} * 12}{women\ aged\ 15{-}44_{y,r}} * 1000\ .$$

The specific month in which the data is collected is *m* (January 1980 and January 1981 considered distinct), the year in which the data is collected is *y*, and the race of the population in question is *r*. This formula scales births annually, and represents the expected annual number of births per thousand women aged 15-44, a commonly used unit in the fertility literature.

[13] SEER counts are based on decennial census estimates, and account for births, deaths, military relocations, and foreign migration; they are unable to account for most domestic migration between census years. The gap between expected population and actual population in census years is retroactively narrowed using an exponential formula, available here (formula at the bottom of page 1): https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/intercensal/intercensal-nat-meth.pdf.

Pollutant Data

I use EPA National Ambient Air Quality Standards (NAAQS) daily annual data files to measure pollution. I aggregate daily observations to the monthly level.[14] In regressions, I restrict the analysis to data from monitors intended to measure the exposure of the local population (8,672 monitors, or 30% of monitors).[15, 16] Monitor placement is likely nonrandom; however, enough of the population live in a county with a monitor (Figure 1) that my estimates provide important information. I use data on particulate matter (both fine particulate matter, or PM 10, and very fine particulate matter, or PM 2.5) and all four criteria gasses (ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2)) on which the EPA collects data.

---

[14] Monthly averages are constructed by taking the geometric mean of all observations for a given population monitor in the specified month, then taking the geometric mean of all population monitors in a county.

[15] Imputed attainment designation uses all EPA-approved monitors to determine attainment status, but the value of the pollutant included in the regression must be from a population monitor.

[16] Other purposes for monitors include "extreme downwind" (identifying routes by which large amounts of pollution are spreading from one region to another through wind patterns); "quality assurance" (making sure measurements are consistent across monitor types and designs); and "source oriented" (if local pollution is generated by a primary organization, isolating that actor). More information can be found here: https://www3.epa.gov/ttnamti1/pamsmain.html; https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/Final%20Handbook%20Document%201_17.pdf.

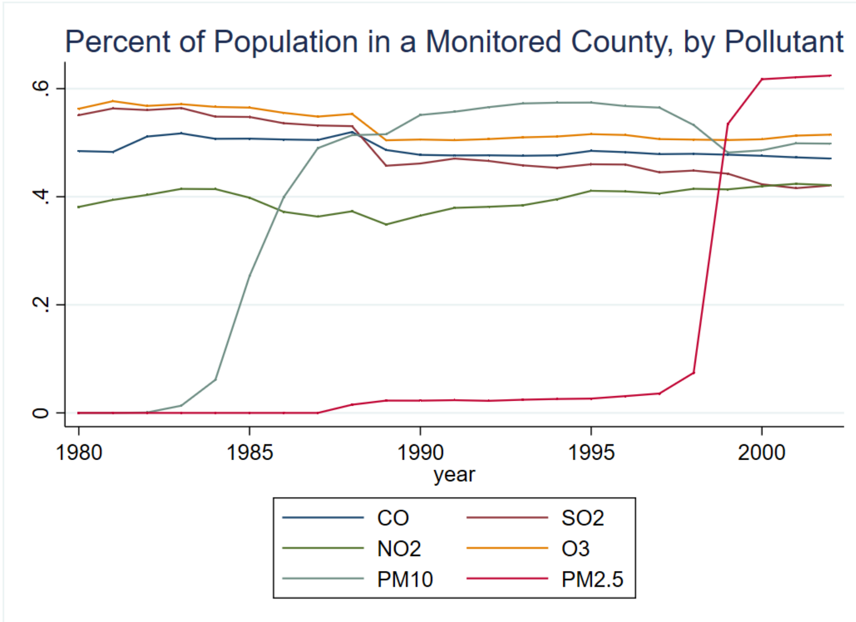

Figure 1: The type of particulate matter observed by the EPA changed over time, from total suspended particulates (TSP) in the 1970s to PM 10 in the 1980s to PM 2.5 in the late 1990s; this is why monitor prevalence changes so dramatically for these pollutants over this time.

Descriptive Statistics

Table 1 provides a short executive summary of key statistical metrics for understanding the distributions of my key variables. In the regressions to follow, I restrict the sample to county-months with over 1,000 women of the population of interest aged 15–44 in order to prevent outlier fertility rates from small counties from obscuring results (further explanation and unrestricted descriptive statistics in Appendix). This means that the samples run in each regression are slightly distinct. See Figures 1 and 2 for distributional summaries of fertility rates and fertility in the dataset by sample.

Figure 3 displays the distribution of pollution levels relative to cutoff values for pollutants whose NAAQS standards changed over time. Particulate matter cutoffs tend to be more binding than ozone cutoffs. Figure 4 displays monthly pollution exposure over time, with darker lines representing higher percentiles. Pollutant levels decline between 1980 and 2002, and seasonal variation becomes slightly less substantial during this period.

Table 1: Descriptive Statistics

	Mean	Standard Deviation	N
O_3, population monitors (ppm)	0.026	0.0106	54,499
PM 10, population monitors (micrograms per m^3)	26.790	12.087	41,192
PM 2.5, population monitors (micrograms per m^3)	13.232	5.209	15,722
SO_2, population monitors (ppb)	7.123	5.636	33,852
NO_2, population monitors (ppb)	18.691	8.669	26,610
CO, population monitors (ppm)	0.886	0.539	31,623
Fertility rate, all	69.070	20.250	376,368
Fertility rate, white	66.664	19.997	369,060
Fertility rate, black	79.802	27.334	139,344

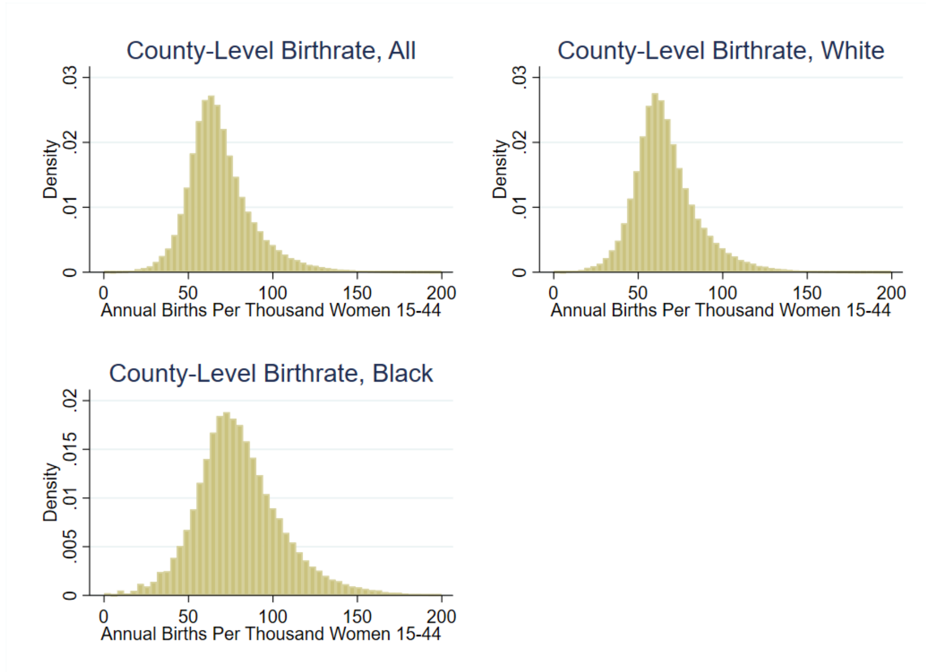

Figure 2: County-Level Fertility Rates, by Race

I limit the range of these plots to fertility rates between 0 and 200 in order to prevent outliers from obscuring the distribution's shape. This limitation means that some observations are not observed in these figures. This omits 0.04% of observations in the pooled histogram, 0.04% in the white histogram, and 0.18% in the black histogram.

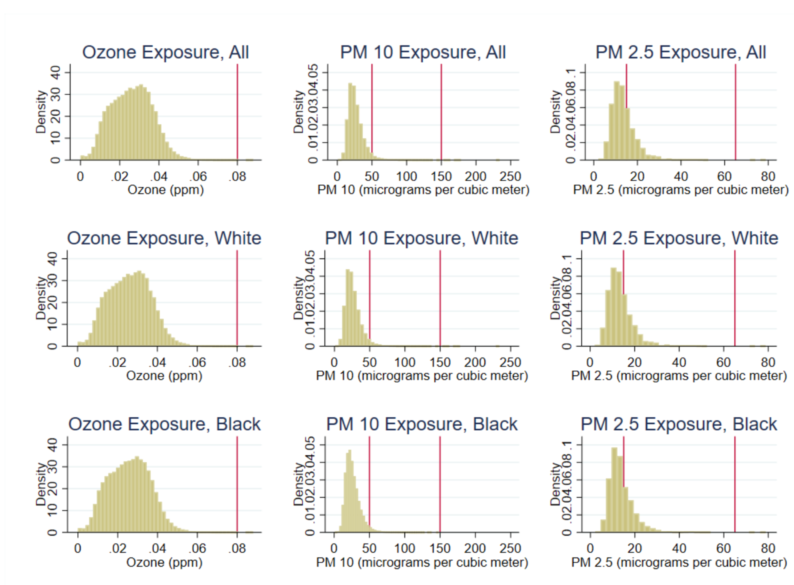

Figure 3: Measured pollutant levels, for monitors with population exposure as an intended purpose

Each row is a sample population, and each column is a pollutant. The red lines represent attainment threshold values. The rightmost line represents some average of maximum values (see Table 2 for exact definitions). On the ozone exposure graph, the rightmost line is off the x axis. The leftmost line generally represents some sort of requirement for the annual mean value.

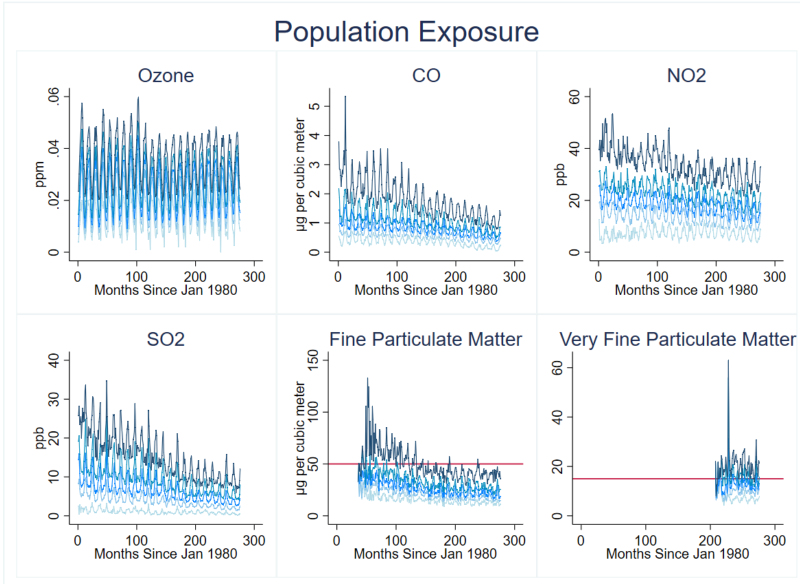

Figure 4: Exposure measured on population monitors, by month. Blue lines represent the 95th, 75th, 50th, 25th, and 5th percentiles in a given month. Red lines represent regulatory thresholds set during the period. The thresholds for pollutants without red lines are so high as to be off the scale; some of these are limits on maximum levels rather than means, and others are simply nonbinding from 1980–2002.

IV. Research Design

OLS Model

This paper considers outcomes exactly nine months after pollution exposure. In the short-term, both women and men are less fertile when exposed to pollution, and women are more likely to experience pregnancy loss (Leiser et al., 2019). I first run an OLS regression of fertility rates on pollution with month and county fixed effects to estimate the association between pollution and fertility, with and without population weights.[17] These effects should absorb variation that is county-specific and time invariant, as well as variation over time that takes place in all counties. However, omitted variable bias is still likely to affect this model. For instance, if changes in unobserved factors are correlated with changes in pollution and changes in fertility, my OLS point estimate would be a biased estimate of the true parameter. Since nonwhite women

[17] I.e., January 1981, February 1981, and January 1982 each have unique indicator values.

are more likely to live in more heavily polluted areas within a county (Gray 2014), I also run regressions using race-specific fertility rates as outcome variables to detect what are likely stronger effects. In conjunction with the minimum population restriction of 1,000 and discussed in my summary statistics, this means that my samples are slightly different depending upon the outcome under consideration.

My OLS baseline specification is of the form:

$$Fertility\ rate_{m+9,c,r} = \alpha + \beta_i\ pollutant_{i,c,m} + \delta_c + \tau_m + X_{c,y}$$

where $Fertility\ rate_{m+9,c,r}$ is the fertility rate in county c during month m for race r; α is a constant; β_i is a unique coefficient for the pollutant included (either ozone, PM 10, or PM 2.5); δ_c are county fixed effects; and τ_m are month fixed effects. $X_{c,y}$ is a vector of county-year controls; this includes percent black population, the unemployment rate, logged county population, and indicators for the distribution of women in the county of the population in question aged 15–44 using 5-year bins. To check the model's strength, I also consider quadratic functional forms, county time trends, and restriction to physically small counties that may be better measured using only one monitor.

IV Model

I hope to discern the degree to which air pollution causes a reduction in fertility given that people can change their behavior to mitigate pollutant exposure. As discussed above, omitted variables are still a very real concern for this model even with the inclusion of time and county fixed effects, especially given the relationship between pollution, fertility, and urban status.

The economics literature studying air pollution makes use of the many natural experiments produced by regulation and technological change.[18] Following Walker (2013) and Chay and Greenstone (2005), who look at housing prices, and Bishop et al. (2018), who examine dementia, I use variation in air pollution induced by the Clean Air Act's National Ambient Air Quality Standards (NAAQS) to estimate the effect of air pollution on fertility (Table 2). The instrument is attainment of a NAAQS standard. The designation of NAAQS standards should have caused a larger reduction of the regulated pollutant for counties in violation of the NAAQS standard (nonattainment counties) than for counties in line with the NAAQS standard (attainment counties), as nonattainment counties worked to get below the standard to avoid facing pecuniary consequences.

In practice, the EPA is conservative in officially designating nonattainment status; more places violate attainment standards than are officially designated nonattainment.[19] However, counties over the NAAQS limit, faced with the threat of official designation, are still incentivized to lower their pollution. Therefore, I impute attainment status based on the EPA's published attainment rules rather than using officially designated attainment status in order to more accurately capture the practical impact of the policy.[20]

[18] For example, Currie and Walker (2011) use address-level data to look at how the sharp reduction of local pollution that came with the introduction of EZ-Pass in New Jersey and Pennsylvania around 2000 affected infant health for pregnant women living near toll areas of major roadways.
[19] Once designated nonattainment, a county must reduce their pollution levels and apply to the EPA for redesignation. If the EPA does not deem their efforts sufficient, the county may face the removal or reduction of federal highway funds (Clean Air Act Amendments, 1970).
[20] Using imputed attainment is common in the literature around pollution, though sometimes it is because the EPA does not have attainment records going back to 1970, when NAAQS standards were first set. Chay and Greenstone (2005) and Walker (2013) both use imputed attainment as their instrument.

Table 2: Pollutant Standards[21]

Pollutant	Enacted	Standard	Designation period
Ozone	Feb 8, 1979	No more than one day per year with max hourly ozone levels above 0.12 ppm	Daily
Ozone	July 18, 1997	Annual fourth-highest daily maximum 8-hr concentration, averaged over 3 years must be less than 0.08ppm	Annual
PM10	July 1, 1987	Over a three-year period, there should not be more than three days with PM10>150 micrograms per cubic meter	Daily
PM10	July 1, 1987	Annual arithmetic mean should not exceed 50 micrograms per cubic meter	Annual
PM2.5	July 18, 1997	The annual 98th percentile, averaged over 3 years, cannot exceed 65	Annual
PM2.5	July 18, 1997	The annual arithmetic mean, averaged over 3 years, cannot exceed 15	Annual

Any of these legislative changes are candidates for instruments that can be used to try to estimate the causal effect of pollution on fertility. In the broadest of terms, for an estimate to be calculated the instrument must affect enough counties to provide both attainment and nonattainment groups of suitable size; this disqualifies using 1997 ozone regulations, as only five counties were nonattainment and contained population monitors. Using an instrumental variable requires four assumptions to be met: monotonicity, relevance, independence, and exclusion. These are discussed below.

a. Monotonicity. This criterion requires that no observations are defiers; that is, no county's level of a pollutant would increase if they were designated nonattainment for that pollutant relative to if they had been designated attainment. One scenario in which this could occur is

[21] See Appendix for specifics on how standards were coded.

if counties very far above permitted levels know they cannot reasonably get to attainment levels, and instead accept any pecuniary consequences and try to attract industry from other counties trying to reduce pollution levels once attainment levels are released. I have no evidence of any behavior of this sort.

b. Relevance. This criterion requires being above the NAAQS standard to be associated with a greater drop in the levels of that pollutant than being below the standard. This is empirically testable in the first stage of a 2SLS regression. A possibility is that monitors are not very good proxies for county pollution, which could cause the instrument to be relatively weak. Bento and Freedman (2014) note that the areas around monitors may experience larger reductions in pollution than the county as a whole. See Appendix for information on the monitoring radius of each monitor; this is a real concern for this project, as counties are usually larger than these radii.

Figure 6 shows average pollution levels in attainment and nonattainment counties, as defined by PM 10 1987 standards and imputed using 1986 pollution levels.

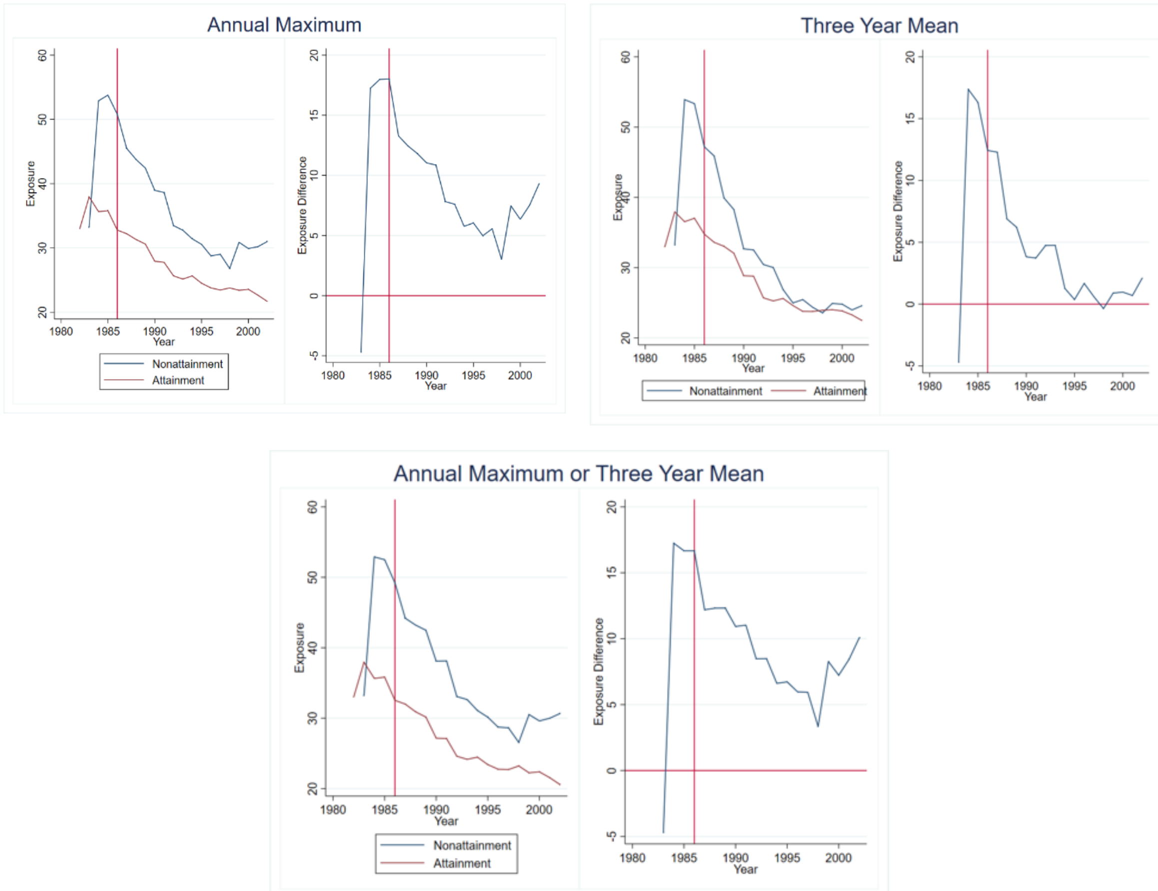

Figure 6: All three instruments show sharp drops after the instrument is introduced. The left panels show annual averages of particulate matter (to prevent seasonal variation from obscuring the pattern), where the blue line represents the counties nonattainment for the instrument in question in 1986 and the red line represents the counties whose values are below the 1987 attainment value in 1986. The right panel shows the difference between annual averages of attainment and nonattainment counties, where the horizontal red line represents no difference between average levels in attainment and nonattainment counties and the vertical red line is 1987, the year in which the PM 10 was regulated. The sharp convergence of annual averages that begins in 1987 for all three instruments supports the relevance criterion.

c. Independence. This criterion requires that, conditional on controls, the instrument as good as randomly assigned. Given the county and year fixed effects included in this model, if *time-static* county-level factors are systematically different in attainment counties versus nonattainment counties, the assumption is not violated, but if *time-variant* county-level factors correlated with fertility rates are systematically different by attainment status the criterion is violated. For example, if nonattainment counties have higher manufacturing concentrations in the years I observe, county fixed effects will absorb this variation and not violate the independence criterion. If nonattainment counties are developing at a different rate than attainment counties, however, this would not be picked up in my model and would violate the independence criterion.

d. Exclusion. This criterion asserts that attainment status only affects fertility through exposure to the regulated pollutant. This is the assumption in the most danger of being unknowingly violated; if a county is trying to get their pollution levels to decrease, the way they do this may impact fertility. For example, a county or state could increase regulation or oversight of industry; this may cause workers to lose their jobs or have lower salaries, and change their desired number of children.

As a way to look at the precision of methods of lowering levels of a pollutant, I examine levels of other pollutants in my dataset based on attainment status for PM 10. If nonattainment counties lower their levels of nonregulated pollutants, it implies counties lower pollution rather bluntly, in ways that may affect fertility rates through other channels. This would threaten the validity of IV estimates. If nonattainment counties raise their levels of nonregulated pollutants, it suggests that they are substituting one type of pollution for another. If, however, pollution rates by attainment status neither converge nor diverge when a different pollution is regulated, then counties are able to regulate pollution in a relatively careful way that plausibly may not affect fertility. Figure 7 displays the average levels of pollutants other than particulate matter over time in PM 10 attainment and nonattainment counties, with one figure per attainment definition. Counties that show up in the figures below had multiple monitor types and thus may be different than differently monitored counties, but this is the strongest indication of the magnitude of possible exclusion restriction violations that I can produce with this data.

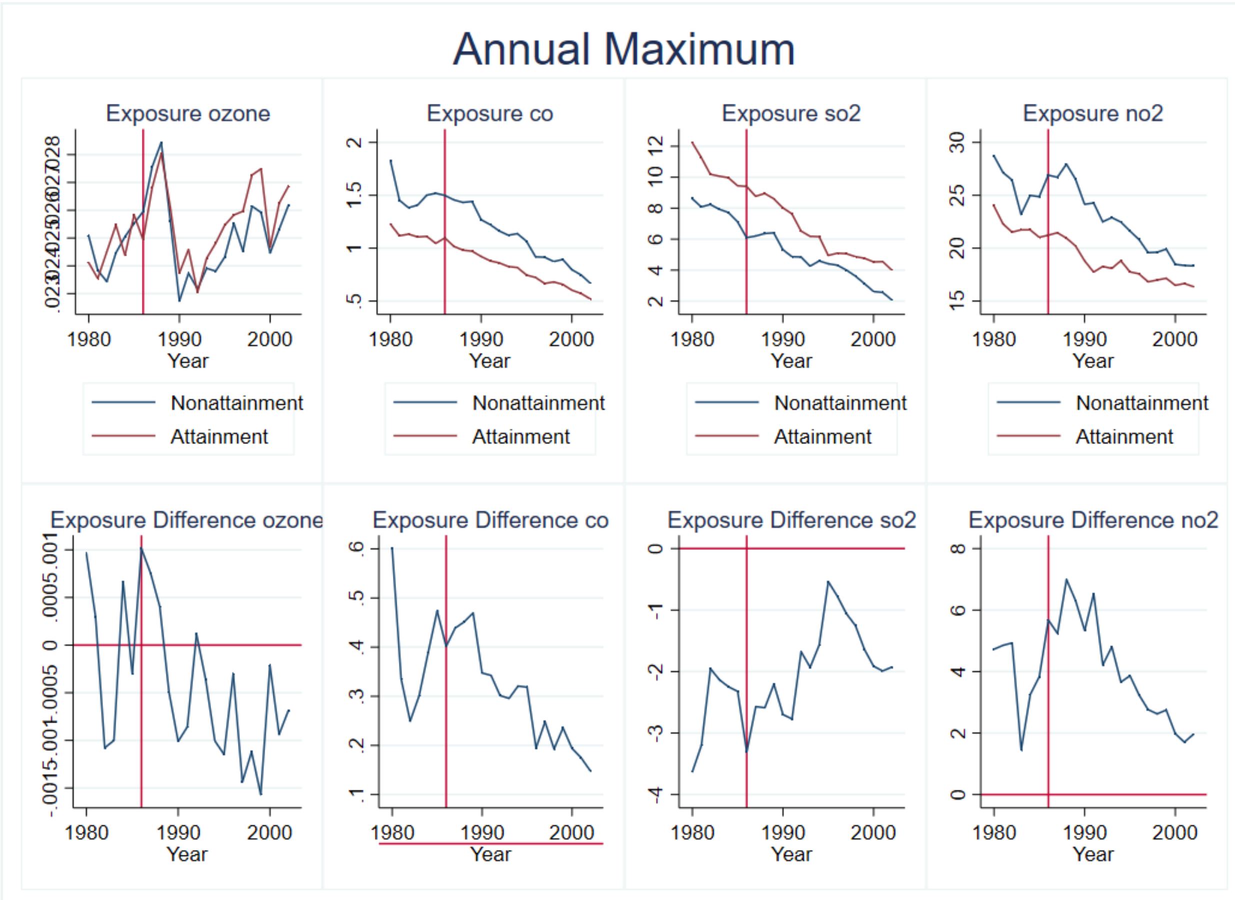

Figure 7. Note: Red vertical lines mark 1987, the year PM 10 standards were set. The first row shows levels of pollution for each type of county; the second row shows the difference between attainment and nonattainment counties in the cell above, with 0 marked horizontally.

For the regulations on annual maximum levels, it appears from Figure 7 that the exclusion restriction is not well met. For all pollutants but ozone, annual levels seem to begin to converge after the standard is set in place. If levels of other pollutants were the only thing changing, these could be controlled for and the instrument would still be valid. However, it seems likely that these convergences are byproducts of larger economic shifts in nonattainment counties, shifts which may affect fertility rates in unobserved ways. If this is happening, an IV estimate will be misleading.

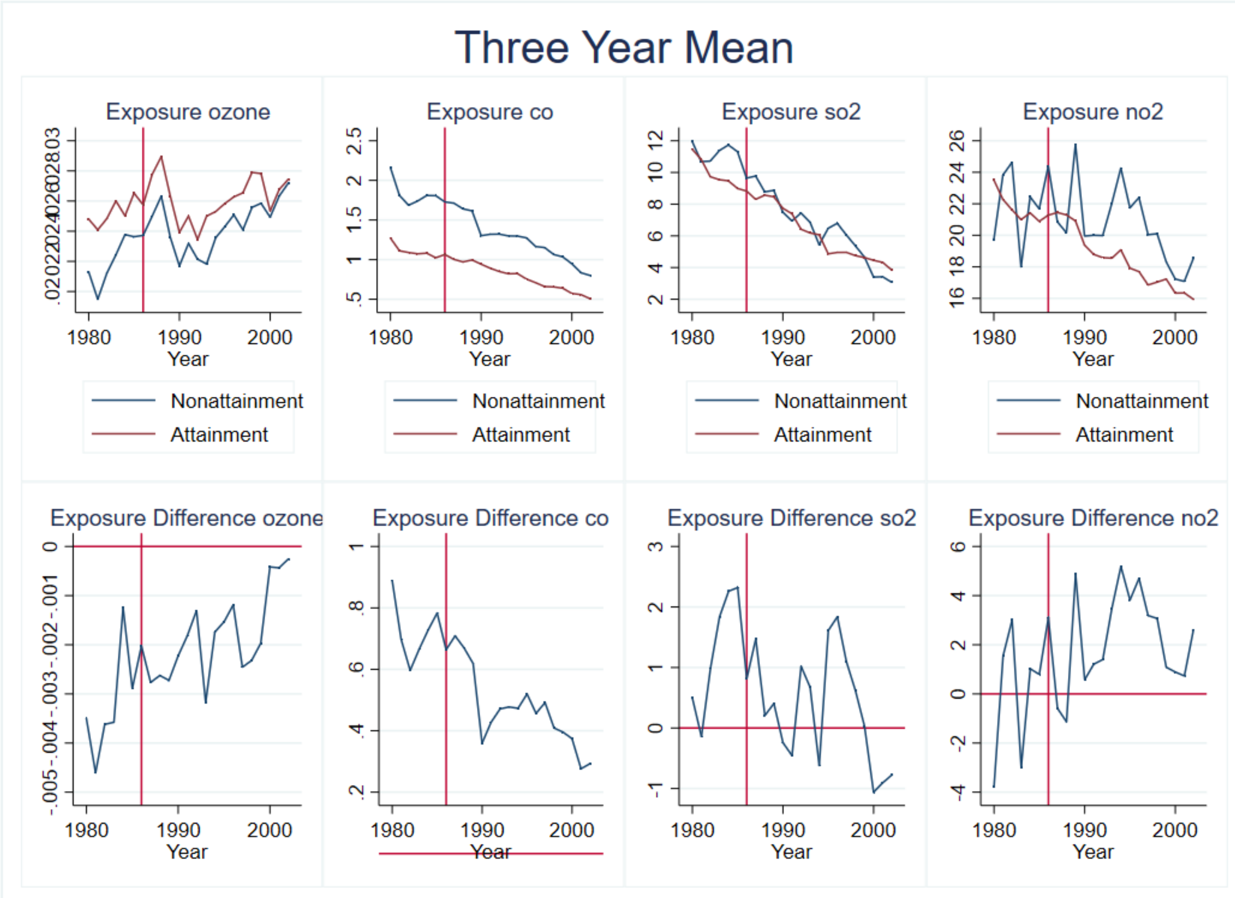

Figure 8. Note: Red vertical lines mark 1987, the year PM 10 standards were set. The first row shows levels of pollution for each type of county; the second row shows the difference between attainment and nonattainment counties in the cell above, with 0 marked horizontally.

For the next attainment designation, the restriction on three-year average levels, Figure 8 shows that nonattainment and attainment counties seem to head towards convergence after the new PM 10 standard is enacted. SO_2 and NO_2 have noisier trends, and it's not clear that attainment status for PM 10 is influential in determining levels of these pollutants. This indicates that this change to the standards is likely not meeting the exclusion restriction.

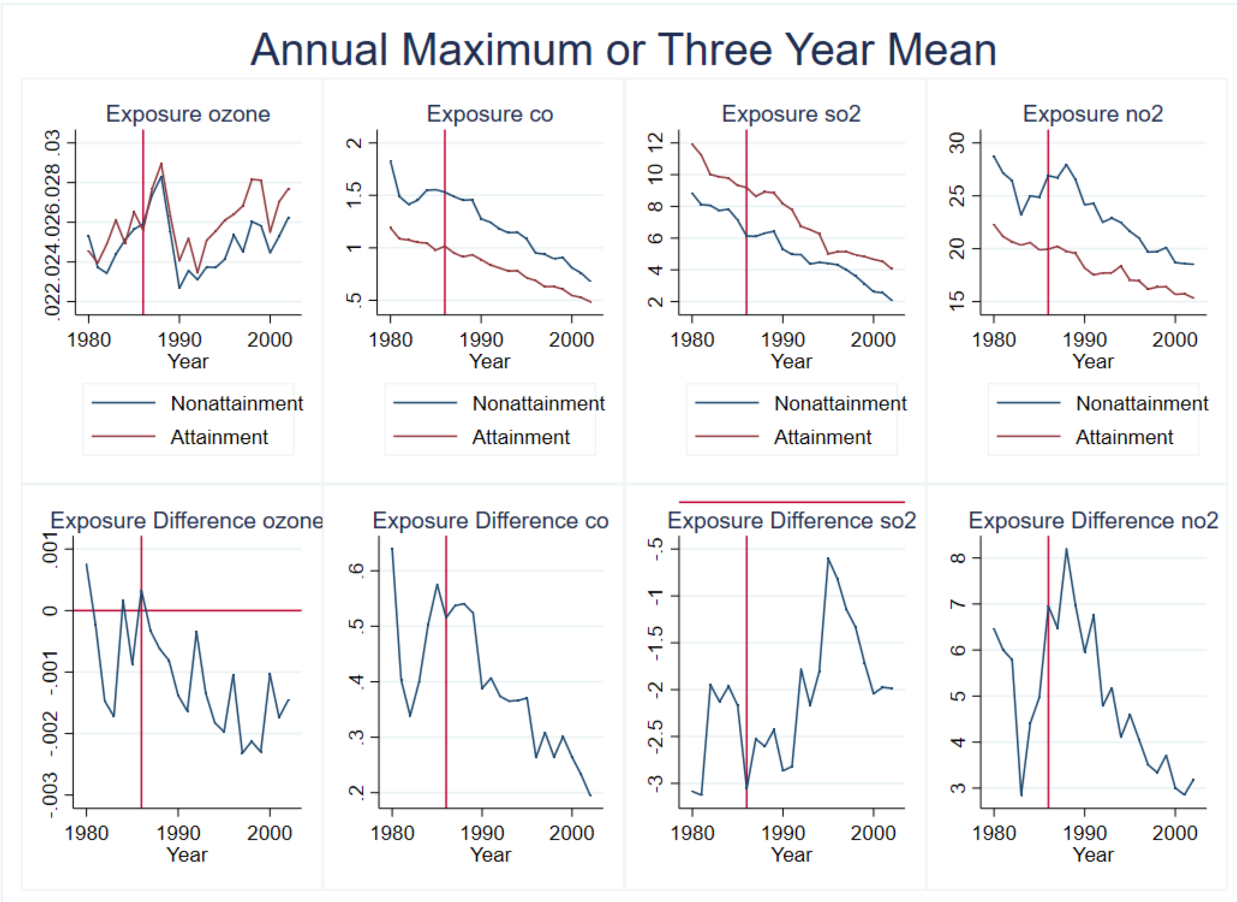

Figure 9. Note: Red vertical lines mark 1987, the year PM 10 standards were set. The first row shows levels of pollution for each type of county; the second row shows the difference between attainment and nonattainment counties in the cell above, with 0 marked horizontally.

For an indicator of whether a county violated either 1987 PM 10 standard, in Figure 9 it appears that attainment status is quite influential in determining levels of other pollutants. This indicates that the exclusion restriction may not be met, and the IV estimate may not be causal.

For the instrumental variables estimate, I compare results aggregated to the year level in two years, one soon before the implementation of the new regulations and one afterwards. The first stage of my IV model estimated using attainment status is specified as:

$$Pollutant\ exposure = \alpha + \beta * attainment\ status + \delta_c + \tau_y + X_{c,y}.$$

This model should isolate the variation in pollutant exposure driven by attainment designation status and my observed covariates. Counties above the attainment threshold are given an incentive to reduce their levels of the pollutant in question that counties below the threshold do not have. In order to meet the relevance criterion, the coefficient on imputed attainment status should be positive.

The second stage of my IV model uses the extra drop in pollution predicted by attainment status to predict fertility rates:

$$\overline{Fertility\ rate} = A_2 + B2\ \ pollut\widehat{ant\ ex}posure\ + \delta_c + \tau_y\ + X_{c,y}$$

If the instrument is to be believed, the coefficient on predicted pollutant exposure in the regression above can be interpreted as the causal effect of a given pollutant on fertility rates. This coefficient is expected to be negative; an increase in a component of pollutant exposure should be associated with decreased fertility.

V. OLS Results

Tables 3 displays my OLS results without population weights, while Table 4 presents results weighted by the number of women 15–44 of the population of interest living in the county. Standard errors are clustered at the county level. Though often statistically insignificant, these results give an indication of some likely weaknesses of an OLS model. I calculate the mean and standard deviation of each pollutant across all county-month observations and normalize the measures such that each pollutant has mean 0 and standard deviation 1. Therefore, the coefficients can be interpreted as the association between a one standard deviation change in levels of the pollutant and fertility outcomes (in my sample).[22] The coefficients in Table 3 can be

[22] This one standard deviation change incorporates all county-years in my sample. This means that the unit change may or may not be a realistic change in pollution levels for a given county.

interpreted as the association between the explanatory variable and the outcome in the average county. Coefficients in Table 4 can be interpreted as the association between the explanatory variable and the outcome in the county in which the average woman lives.

For example, consider Panel A, Specification (4) in both of these tables. In Table 3, if a monitored county were selected at random, a one standard deviation increase CO is expected to be associated with a 0.089 decease in that county's fertility rate, measured in births per thousand women. In Table 4, if a woman aged 15–44 in a PM CO-monitored county were selected at random, a one standard deviation increase in CO is expected to be associated with a 0.167 unit decrease in her county's fertility rate (a 0.2% decrease from the mean fertility rate).

Both tables exhibit similar coefficient behavior across panels. With the exception of sulfur dioxide, coefficients on pollution tend to have a positive sign and relatively small magnitude. As the science stands now, there is no plausible biological pathway for increased pollution to directly cause higher fertility, so I interpret these positive coefficients to mean that omitted variables are likely biasing my estimates even with fixed effects to absorb cross-county time-invariant differences and nationally consistent, time-variant changes. Results for black women tend to be more negative than results for the pooled sample or white women, but across pollutants the main takeaway from these tables is the high variability of these estimates.

These results are highly dependent upon the specification. Adding county time trends, quadratic pollutant functional forms, and county size cutoffs all change the resulting coefficients quite a bit, affecting both magnitude and sign. This lack of robustness to specification suggests that OLS is a poor identification strategy here.

Table 3: OLS by Race Regression Results

Specification:	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Women						
Ozone	0.412***					
	(0.133)					
PM 10		0.031				
		(0.088)				
PM 2.5			0.073			
			(0.089)			
CO				-0.089		
				(0.154)		
NO_2					0.285	
					(0.198)	
SO_2						-0.239
						(0.154)
R squared	0.813	0.824	0.872	0.841	0.846	0.784
n	50,606	37,390	12,276	29,473	24,661	31,564
Panel B: White Women						
Ozone	0.265**					
	(0.134)					
PM 10		0.072				
		(0.096)				
PM 2.5			0.110			
			(0.096)			
CO				0.148		
				(0.195)		
NO_2					0.385*	
					(0.230)	
SO_2						-0.250
						(0.170)
R squared	0.815	0.835	0.873	0.841	0.850	0.783
n	50,605	37,390	12,276	29,473	24,661	31,564
Panel C: Black Women						
Ozone	1.307***					
	(0.301)					
PM 10		-0.231				
		(0.178)				
PM 2.5			0.207			
			(0.172)			
CO				-0.402		
				(0.346)		
NO_2					-0.174	

					(0.316)	
SO_2						-0.572
						(0.392)
R squared	0.501	0.541	0.534	25,954	21,148	26,490
n	41,987	30,548	10,423	0.548	0.562	0.520
County FE	X	X	X	X	X	X
Year-Month FE	X	X	X	X	X	X
County-Level Controls	X	X	X	X	X	X

Notes: *p<0.1; **p<0.05; ***p<0.01. Observations are at the county-month level and are restricted to counties with over 1000 women of the population of interest and monitors that are placed specifically to measure population exposure. Pollutant levels are standardized to have a mean of 0 and standard deviation of 1, so coefficients are interpreted as the change in fertility rates associated with a one sample standard deviation increase in a given pollutant. This standardization is done before regressions are run, so the unit change for a given pollutant is consistent across panels and tables even though the counties with outcome observations may differ. County covariates are percent black, log population, and 5-year age bins for the distribution of the female population 15–44. Standard errors are clustered at the county level.

Table 4: OLS by Race Regression Results, Population Weighted

Specification:	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Women						
Ozone	0.323*					
	(0.192)					
PM 10		0.195*				
		(0.103)				
PM 2.5			0.128			
			(0.078)			
CO				-0.169		
				(0.211)		
NO_2					0.130	
					(0.213)	
SO_2						-0.899**
						(0.425)
R squared	0.848	0.856	0.901	0.848	0.8661	0.852
n	50,606	37,390	12,276	29,473	24,661	31,564
Panel B: White Women						
Ozone	0.129					
	(0.212)					
PM 10		0.270**				
		(0.110)				
PM 2.5			0.178			
			(0.095)			

CO				0.066		
				(0.284)		
NO_2					0.009	
					(0.337)	
SO_2						-1.238**
						(0.566)
R squared	0.861	0.884	0.916	0.870	0.876	0.867
n	50,605	37,390	12,276	29,473	24,661	31,564
Panel C: Black Women						
Ozone	1.461***					
	(0.314)					
PM 10		-0.029				
		(0.176)				
PM 2.5			0.056			
			(0.132)			
CO				-0.367		
				(0.335)		
NO_2					0.347	
					(0.249)	
SO_2						0.083
						(0.409)
R squared	0.676	0.740	0.700	0.731	0.749	0.718
n	41,987	30,548	10,423	25,954	21,148	26,490
County FE	X	X	X	X	X	X
Year-Month FE	X	X	X	X	X	X
County-Level Controls	X	X	X	X	X	X

Notes: *p<0.1; **p<0.05; ***p<0.01. Observations are at the county-month level and are restricted to counties with over 1000 women of the population of interest and monitors that are placed specifically to measure population exposure. Pollutant levels are standardized to have a mean of 0 and standard deviation of 1, so coefficients are interpreted as the change in fertility rates associated with a one sample standard deviation increase in a given pollutant. This standardization is done before regressions are run, so the unit change for a given pollutant is consistent across panels and tables even though the counties with outcome observations may differ. County covariates are percent black, log population, and 5-year age bins for the distribution of the female population 15–44. Standard errors are clustered at the county level.

VI. IV Results

In light of the likely biased OLS results reported above, I provide IV estimates of the effect of PM 10 on fertility. Table 5 contains first stage results, and Table 6 contains second stage results. Between limitations on the number of PM 10 monitors in 1986 and the restriction of births data after 1988, my IV regressions have only about 130 counties. With a valid IV, these coefficients can be interpreted as a causal estimate; this IV has a few weaknesses that may prevent that interpretation, but these coefficients are still preferred to the OLS tables above. For simplicity and ease of interpretation, these estimates are calculated by comparing counties using only two time periods, one before the 1987 designation and one after. I aggregate to the year level to prevent seasonality from affecting these estimates. Below are comparisons of 1986, the last year before redesignation in 1987 and thus the closest measure of pre-designation pollution levels, to a few years after the designation. Notably, the first stage is weak for all estimates; ideally, this F statistic would be at least 10 for an instrument to be considered strong. This is an issue with the relevance criteria and is likely do to the relatively small sample size when comparing across only two time periods for counties with monitors in both.

The panels in Table 5 and Table 6 use 1986 as the first year and either 1988, 1990, 1992, 1994, or 1996 as the second year. This aims to capture the timing of the legislation's impact and balance increased noise with increased efficacy of the standard over time. The first stage gets stronger each year, with a levelling-off in 1994. This implies that it takes a few years for counties to lower their pollution levels after pollution regulations are announced. Second stage results are harder to interpret, which is at least partially due to a weak first stage, and there is not a clear trend in the results. Specifications using 1990 and 1992 as final years are the most in line with expected signs and magnitudes. No other NAAQS standards were redesignated between 1986 and 1996.

Though this issue plagues all three legislative instruments, the first and third columns have much stronger first stages and estimates closer to statistical significance. In the initial years after the standard is released, the estimated impact of a given PM 10 reduction is larger than the impact of an identical reduction later on. This suggests that the relationship between pollution and fertility may be nonlinear. It could also be due to the use of imputed attainment rather than official attainment; it seems likely that counties above the standard that are not officially designated nonattainment perform the bulk of their reduction soon after the release of the standard. It is surprising that the coefficients on black fertility rates are so different than those on the white and pooled samples, but not in a systematic way. If it is not statistical noise, this could be because economic consequences violate the exclusion restriction and hit this population harder, or because, if population monitors are more likely to be placed in white areas of a county, reductions in pollution occurred in areas near monitors and did not benefit the full county.

Columns 4–6 are weighted by population. If the causal effect of PM 10 on fertility is the same across places, the coefficients in these columns would be expected to be identical to those in columns 1–3. Almost none of these values are statistically significant, although this could be statistical noise. This could also indicate treatment effect heterogeneity. For instance, if a county where a high proportion of residents spend substantial time outdoors had a decline in pollution, this may have a larger causal effect on fertility than an identical decline in a county where people spend more time indoors because residents are less exposed to ambient outdoor levels.

Table 5: IV Regression Results, First Stage

Outcome: Race-Specific Fertility Rate	(1) Annual Max	(2) Three Year Mean	(3) Annual Max or Three Year Mean	(4) Annual Max, Population Weighted	(5) Three Year Mean, Population Weighted	(6) Annual Max or Three Year Mean, Population Weighted
Before = 1986, After = 1988						
PM 10, All	0.467	-0.070	0.369	0.708	-0.186	0.574
	(0.351)	(0.445)	(0.355)	(0.276)	(0.470)	(0.314)
First Stage F	1.769	0.026	1.082	6.55	0.160	3.349
PM 10, White	0.467	-0.043	0.374	0.622**	-0.083	0.490
	(0.357)	(0.440)	(0.357)	(0.270)	(0.442)	(0.302)
First Stage F	1.716	0.010	1.103	5.290	0.036	2.657
PM 10, Black	0.621	0.156	0.490	0.618	-0.258	0.602
	(0.455)	(0.799)	(0.490)	(0.445)	(0.604)	(0.465)
First Stage F	1.850	0.040	1.000	1.932	0.185	1.664
Before = 1986, After = 1990						
PM 10, All	0.762**	0.476	0.713*	0.838***	0.260	0.756*
	(0.336)	(0.543)	(0.379)	(0.280)	(0.567)	(0.329)
First Stage F	5.153	0.774	3.534	9.000	0.212	5.290
PM 10, White	0.721**	0.517	0.683*	0.761**	0.366	0.689**
	(0.321)	(0.526)	(0.359)	(2.89)	(0.573)	(0.329)
First Stage F	5.063	0.960	3.648	6.97	0.410	4.410
PM 10, Black	0.713	0.450	0.645	1.043*	0.303	1.026*
	(0.451)	(0.926)	(0.481)	(0.529)	(0.785)	(0.603)
First Stage F	2.496	0.240	1.796	3.881	0.152	2.890
Before = 1986, After = 1992						
PM 10, All	0.812**	0.380	0.741*	0.957**	0.041	0.846*
	(0.374)	(0.572)	(0.400)	(0.386)	(0.779)	(0.444)
First Stage F	4.709	0.449	3.423	4.709	0.003	3.648
PM 10, White	0.790**	0.436	0.751*	0.884**	0.094	0.781*
	(0.349)	(0.564)	(0.378)	(0.406)	(0.769)	(0.446)
First Stage F	5.108	0.593	3.960	4.752	0.014	3.063
PM 10, Black	0.844*	0.446	0.772*	1.226**	0.287	1.194**
	(0.426)	(0.780)	(0.433)	(0.497)	(0.666)	(0.571)
First Stage F	3.920	0.325	3.168	6.052	0.185	4.368

Before = 1986, After = 1994						
PM 10, All	1.061***	0.532	1.006**	1.108***	0.288	1.008**
	(0.396)	(0.670)	(0.428)	(0.379)	(0.683)	(0.441)
First Stage F	7.182	0.640	5.523	8.58	0.176	5.244
PM 10, White	1.039***	0.650	0.997**	1.094**	0.378	0.962**
	(0.369)	(0.623)	(0.392)	(0.418)	(0.656)	(0.452)
First Stage F	7.952	1.160	6.45	6.864	0.336	4.537
PM 10, Black	1.134**	0.680	1.063*	1.082***	0.308	1.063**
	(0.545)	(1.084)	(0.608)	(0.389)	(0.890)	(0.449)
First Stage F	4.326	0.397	3.063	7.728	0.123	5.617
Before = 1986, After = 1996						
PM 10, All	1.143**	0.422	1.014*	1.085***	0.299	0.926**
	(0.514)	(0.801)	(0.542)	(0.363)	(0.708)	(0.458)
First Stage F	4.973	0.281	3.497	8.94	0.176	4.08
PM 10, White	1.112**	0.516	0.990*	1.041**	0.375	0.845*
	(0.491)	(0.729)	(0.520)	(0.402)	(0.651)	(0.498)
First Stage F	5.108	0.504	3.610	6.708	0.336	2.890
PM 10, Black	1.025**	0.522	0.907*	1.012***	0.707	0.999***
	(0.457)	(1.050)	(0.516)	(0.259)	(0.708)	(0.306)
First Stage F	5.018	0.250	3.098	15.366	1.000	10.628
County FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
County-Level Controls	X	X	X	X	X	X
Population-Weighted				X	X	X

Notes: *p<0.1; **p<0.05; ***p<0.01. Observations are at the county-month level and are restricted to counties with over 1000 women of the population of interest and monitors that are placed specifically to measure population exposure. Pollutant levels are standardized to have a mean of 0 and standard deviation of 1, so coefficients are interpreted as the change in fertility rates associated with a one sample standard deviation increase in a given pollutant. This standardization is done before regressions are run, so the unit change for a given pollutant is consistent across panels and tables even though the counties with outcome observations may differ. County covariates are percent black, log population, and 5-year age bins for the distribution of the female population 15–44. Standard errors are clustered at the county level.

Table 6: IV Regression Results, Second Stage

Outcome: Race-Specific Fertility Rate	(1) Annual Max	(2) Three Year Mean	(3) Annual Max or Three Year Mean	(4) Annual Max, Population Weighted	(5) Three Year Mean, Population Weighted	(6) Annual Max or Three Year Mean, Population Weighted
Before = 1986, After = 1988						
PM 10, All	-6.671	0.136	-5.832	-0.475	1.580	-0.025
	(4.485)	(21.375)	(4.784)	(1.124)	(6.704)	(1.333)
First Stage F	1.769	0.026	1.082	6.55	0.160	3.349
PM 10, White	-7.336	-1.064	-6.501	-1.190	2.480	-0.819
	(4.748)	(34.595)	(4.948)	(1.594)	(16.301)	(1.915)
First Stage F	1.716	0.010	1.103	5.290	0.036	2.657
PM 10, Black	8.453	0.156	10.204	0.155	-1.043	0.203
	(5.421)	(0.799)	(7.3112)	(2.209)	(5.115)	(2.234)
First Stage F	1.850	0.040	1.000	1.932	0.185	1.664
Before = 1986, After = 1990						
PM 10, All	-2.304	-0.368	-1.682	-1.274	0.079	-0.596
	(1.643)	(3.298)	(1.619)	(1.323)	(6.978)	(1.469)
First Stage F	5.153	0.774	3.534	9.000	0.212	5.290
PM 10, White	-2.686	0.767	-2.103	-2.391	-0.262	-1.732
	(1.870)	(2.951)	(1.837)	(1.764)	(4.939)	(2.017)
First Stage F	5.063	0.960	3.648	6.97	0.410	4.410
PM 10, Black	-0.681	0.128	1.381	-0.337	-13.674	-0.267
	(2.057)	(6.063)	(2.695)	(2.002)	(15.572)	(1.993)
First Stage F	2.496	0.240	1.796	3.881	0.152	2.890
Before = 1986, After = 1992						
PM 10, All	-1.453	-0.116	-1.014	0.560	-0.220	1.155
	(1.709)	(4.567)	(1.844)	(1.408)	(48.153)	(1.679)
First Stage F	4.709	0.449	3.423	4.709	0.003	3.648
PM 10, White	-2.306	1.093	-1.647	-0.578	4.469	0.173
	(1.914)	(4.129)	(1.925)	(1.612)	(26.252)	(2.010)
First Stage F	5.108	0.593	3.960	4.752	0.014	3.063
PM 10, Black	-1.109	-8.059	-0.901	3.194**	-14.436	3.205**
	(2.803)	(7.429)	(2.864)	(1.594)	(16.876)	(1.594)
First Stage F	3.920	0.325	3.168	6.052	0.185	4.368

Before = 1986, After = 1994						
PM 10, All	0.607	-0.422	0.632	1.843	-0.827	2.080
	(1.280)	(3.542)	(1.357)	(1.218)	(6.272)	(1.343)
First Stage F	7.182	0.640	5.523	8.58	0.176	5.244
PM 10, White	0.984	1.392	0.391	1.366	4.543	2.035
	(1.272)	(2.893)	(1.290)	(1.167)	(5.848)	(1.452)
First Stage F	7.952	1.160	6.45	6.864	0.336	4.537
PM 10, Black	-3.448	-8.434	-3.188	1.578	-24.649	1.675
	(3.163)	(8.958)	(3.269)	(2.256)	(42.162)	(2.265)
First Stage F	4.326	0.397	3.063	7.728	0.123	5.617
Before = 1986, After = 1996						
PM 10, All	0.306	-0.808	0.309	0.710	0.058	0.958
	(1.097)	(3.274)	(1.257)	(1.171)	(5.711)	(1.480)
First Stage F	4.973	0.281	3.497	8.94	0.176	4.08
PM 10, White	-0.454	1.326	-0.107	-0.059	5.285	0.845
	(1.139)	(2.784)	(1.243)	(1.181)	(6.807)	(0.498)
First Stage F	5.108	0.504	3.610	6.708	0.336	2.890
PM 10, Black	-2.602	-4.925	-2.101	1.380	-8.022	1.469
	(3.743)	(11.137)	(4.006)	(1.837)	(7.723)	(1.845)
First Stage F	5.018	0.250	3.098	15.366	1.000	10.628
County FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
County-Level Controls	X	X	X	X	X	X
Population-Weighted				X	X	X

Notes: *p<0.1; **p<0.05; ***p<0.01. Observations are at the county-month level and are restricted to counties with over 1000 women of the population of interest and monitors that are placed specifically to measure population exposure. Pollutant levels are standardized to have a mean of 0 and standard deviation of 1, so coefficients are interpreted as the change in fertility rates associated with a one sample standard deviation increase in a given pollutant. This standardization is done before regressions are run, so the unit change for a given pollutant is consistent across panels and tables even though the counties with outcome observations may differ. County covariates are percent black, log population, and 5-year age bins for the distribution of the female population 15–44. Standard errors are clustered at the county level.

VII. Conclusion

Estimates of changes to fertility are important, and reproductive, germ-line consequences are an overlooked effect of pollution. Using an instrumental variables identification strategy, I find a one standard deviation increase in fine particulate matter to be associated with a 1–2 unit decrease in county-level fertility rates, though this is not statistically significant. A weakness of this analysis is the sporadic nature of pollution monitors for the early years of the sample. In the future, a researcher with access to NCHS restricted natality data could conduct an analysis with more consistent pollution data to strengthen these results. Municipality-level analysis for a smaller region of the country could also help pin down these parameters. The NAAQS legislative instrument is worth further examination; though it's used in the literature fairly frequently, possible issues with the exclusion restriction raised when standards affect pollutants they do not directly regulate indicate that its theoretical basis may be weaker than it seems at first glance.

Appendix

NCHS Natality Data

NCHS natality data contains data on the child's race, the mother's race, and the father's race, though which variables are included changes by year. Child's race is not available after 1993, and father's race, though available in all years, is often missing. Mother's race is the most consistently measured race variable, so this is the variable I use to determine race. Pre-1993 it can be coded as missing, while post-1993 it must be filled out.

St. Clair (FIPS 01115) and Shelby (FIPS 01117) counties, both in Alabama, are dropped from the sample, as their fertility rates change dramatically while their populations remain constant in what can only be explained by a coding error.

NCHS only reports the number of births in a county; small counties may not experience any births in a given month, and this lack of births is a meaningful zero. However, some larger counties also are missing births for a month (about 0.6% of the sample in the years before 1989); it's unlikely that there were truly no births in the county, and more likely that the records were somehow distorted. To address this, zeroes are imputed when the number of births in the previous month in that county was 20 or fewer. If the number of births in the previous month was over 20, the births value is left missing. For the first month, the same rule applies but births are compared to the second month because there is no preceding month.

Back to Section III

Population Cutoffs

The standard deviation of the fertility rate variable is quite large for women of color in the unrestricted sample. For example, Towns County, GA (FIPS 13281) is one of the small counties pulling this number up; it had 1,217 female residents aged 15–44 in 1980. Of these women, 1,213 were white, and one was black. There were two births to black mothers (likely twins, though one or both could have been born to women younger than 15 or older than 44) during August, causing the black fertility rate assigned to this county observation to be 24,000 children per thousand women. Especially for women of color, this small-sample ratio issue is not uncommon. Descriptive statistics in the unrestricted sample are below.

Table 1: Descriptive Statistics: Unrestricted Sample

Variable	Mean (Standard Deviation)	Population-Weighted Mean (Standard Deviation)	n
All Women Fertility rate	70.1 (24.7)	66.8 (12.7)	407,413
White Fertility rate	68.2 (24.9)	64.9 (14.0)	405,005
Black Fertility rate	75.2 (358.0)	74.3 (1,16.2)	359,773
All Births	182.2 (552.0)	1,690.8 (3,007.9)	409,381
White Births	100.4 (330.9)	1,122.7 (2,205.1)	352,995
Black Births	22.1 (108.4)	305.1 (583.0)	352,995
All Fertile Population	86,770.6 (30,409.3)	278,002.0 (454,990.3)	448,813
White Fertile Population	65,032.3 (24,592.3)	207,943.2 (334,444.2)	448,813

Black Fertile Population	18,456.5 (4,412.9)	47,178.0 (81,411.8)	448,813
Ozone (ppm)	0.026 (0.011)	0.025 (0.011)	54,499
PM 10 (micrograms per cubic meter)	26.79 (12.09)	28.79 (11.94)	41,192
PM 2.5 (micrograms per cubic meter)			

Back to Section III

Monitor Data

In this paper, readings from a small number of monitors is often intended to proxy for pollutant exposure for women living over a large area of land. Below is a summary of the land area for which monitors are intended to proxy. Larger distances tend to be located in rural areas, since there is less development and local fluctuation in pollutant levels. Restricting pollution data to monitors intended to measure population levels of exposure may mitigate this somewhat.

Ozone Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	15	0.65%
100 m – 500 m	38	1.65%
500 m – 4 km	955	41.56%
4 km – 50 km	857	37.29%
50 km – hundreds km	433	18.84%

PM10 Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	84	2.45%
100 m – 500 m	412	11.99%
500 m – 4 km	2,578	75.05%
4 km – 50 km	218	6.35%
50 km – hundreds km	143	4.16%

PM 2.5 Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	91 3794	2.40%
100 m – 500 m	131	3.45%
500 m – 4 km	2,607	68.71%
4 km – 50 km	518	13.65%
50 km – hundreds km	447	11.78%

NO_2 Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	76	6.64%
100 m – 500 m	76	6.64%
500 m – 4 km	625	54.63%
4 km – 50 km	260	22.73%
50 km – hundreds km	107	9.35%

SO_2 Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	29	1.28%
100 m – 500 m	133	5.87%
500 m – 4 km	1,708	75.38%
4 km – 50 km	268	11.83%
50 km – hundreds km	128	5.65%

CO Monitors: Intended Radius Breakdown

Distance	Count	Percent
0 m – 100 m	291	29.63%
100 m – 500 m	137	13.95%
500 m – 4 km	452	46.03%
4 km – 50 km	60	6.11%
50 km – hundreds km	42	4.28%

Back to Section IV

Attainment Imputation

Attainment designation was often complex (see Table 2), and my imputation required small idiosyncratic judgements. Here I explain the specifics of how each attainment designation was coded.

Ozone

i. Applies to each day until July 18, 1997. If more than one day in the preceding 365 exceeded 0.12 ppm of ozone, the day was designated nonattainment. If a month had one or more nonattainment days, it was designated nonattainment. The cutoff of this definition affects only a small number of observations; 714 observations have between 1 and 15 nonattainment days in a month, while 181,116 have over 15 nonattainment days in a month.

ii. Applies to all months after and including July 1997. Each calendar year, the fourth-highest value was collected and averaged with the fourth-highest daily values from the previous two calendar years. If this value exceeds 0.08 ppm, all 12 months are coded nonattainment for that calendar year (except months before July 1997, which are coded missing). Technically, this measure should be calculated using 8-hour data instead of daily data; my calculation is therefore conservative, and the counties coded nonattainment are the most severe cases.

PM 10

i. Applies to each day after and including July 1, 1987. If more than three days in the preceding 1095 days exceed 150 micrograms per cubic meter of ozone, the day is designated nonattainment by one metric. If a month has one or more nonattainment days, it is designated nonattainment by this metric. The cutoff of this definition affects only a small number of observations; 127 observations have between 1 and 15 nonattainment days in a month, while 160,262 have over 15 nonattainment days in a month.

ii. Applies to each month after and including July 1987. If the mean of all collected measurements in a calendar year exceeded 50 micrograms per cubic meter, each month in that year is designated nonattainment (except months preceding July 1987, which is left missing).

iii. Importantly, though PM 10 was not regulated before 1987, Total Suspended Particulates (TSP), an earlier iteration of particulate measurement that included larger particles, was regulated, so the extent to which changes in PM legislation are binding is unclear. TSP data is not provided online, so I am unable to verify the practical relevance of this standard beyond inferential speculation on the first stage coefficient of this as a legislative instrument.

PM 2.5

i. Applies to all months after and including July 1997. Each calendar year, the value in the 98th percentile was collected and averaged with the 98th percentile values from the previous two calendar years. In cases where not enough values were collected to have an even 98th percentile, this was replaced with the 97th (269 county-years), 96th (27 county-years), or 95th (11 county-years) percentiles as needed. If this value exceeds 0.08 ppm, all 12 months are coded nonattainment for that calendar year (except months before July 1997, which are coded missing).

ii. Applies to all months after and including July 1997. Each calendar year, mean of all values collected in that calendar year is averaged with the means collected in the two previous calendar years. If this value exceeds 15 micrograms per cubic meter, all 12 months are coded nonattainment for that calendar year (except months before July 1997, which are coded missing).

iii. Both TSP and PM 10 were regulated at this point, so PM 2.5 regulations may not be binding in all places.

Back to Table 2

References:

Agency for Toxic Substances and Disease Registry. "Public Health Statement Sulfur Dioxide." December 1998. https://www.atsdr.cdc.gov/ToxProfiles/tp116-c1-b.pdf.

Amelar, Richard D., Lawrence Dubin, and Cy Dchoenfeld. "Sperm Motility." *Fertility and Sterility* 34:3. 1980.

American Industrial Hygiene Association. "Odor Thresholds for Chemicals with Established Health Standards, 2nd Edition." 2013. https://www.pdo.co.om/hseforcontractors/Health/Documents/HRAs/ODOR%20THRESHOLDS.pdf.

Bento, Antonio and Matthew Freedman. "Who Benefits from Environmental Regulation? Evidence from the Clean Air Act Amendments. *Review of Economics and Statistics* 97:3, 2014.

Bishop, Kelly C., Jonathan D. Ketcham, and Nicolai V. Kuminoff. "Hazed and Confused: The Effect of Air Pollution on Dementia." NBER Working Paper 24970, 2018.

Boguski, Terrie K. "Understanding Units of Measurment." Environmental Science and Technology Briefs for Citizens, Center for Hazardous Substance Research. October 2006. https://cfpub.epa.gov/ncer_abstracts/index.cfm/fuseaction/display.files/fileid/14285.

Carre, Julie, Nicolas Gatimel, Jessika Moreau, Jean Parinaud, and Roger Leandri. "Does Air Pollution Play a Role in Infertility? A Systematic Review." *Environmental Health* 16:82, 2017.

Centers for Disease Control and Prevention. "Morbidity and Mortality Weekly Report." June 23, 2017. https://www.cdc.gov/mmwr/volumes/66/ss/ss6613a1.htm.

Chay, Kenneth Y. and Michael Greenstone. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy* 113:2, 2005.

Choe, SA, YB Jun, WS Lee, TK Yoon, SY Kim. "Association Between Ambient Air Pollution and Pregnancy Rate in Women who Underwent IVF." *Human Reproduction* 33(6). June 2018. https://academic.oup.com/humrep/article/33/6/1071/4962139.

Clay, Karen, Margarita Portnykh, and Edson Severini. "Toxic Truth: Lead and Fertility." NBER Working Paper 24607. May 2018.

Clean Air Act Amendments of 1970. Public Law 91-604. https://www.gpo.gov/fdsys/pkg/STATUTE-84/pdf/STATUTE-84-Pg1676.pdf.

Currie, Janet, Matthew Neidell, and Juhannes Schmeider. "Air Pollution and Infant Health: Lessons From New Jersey." *Journal of Health Economics* 28(3). May 2009. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2727943/.

Currie, Janet and Reed Walker. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics 3*, 2011.

Dales, Robert, Richard T. Burnett, Marc Smith-Doiron, David M. Stieb, and Jeffrey R. Brook. "Air Pollution and Sudden Infant Death Syndrome." *Pediatrics* 113:6. 2004.

Environmental Protection Agency. "Nitrogen Dioxide Trends." Accessed April 11, 2019. https://www.epa.gov/air-trends/nitrogen-dioxide-trends.

Environmental Protection Agency. "Photochemical Assessment Stations (PAMS)". Accessed April 11, 2019. https://www3.epa.gov/ttnamti1/pamsmain.html.

Environmental Protection Agency. "Quality Assurance Handbook for Air Pollution Measurement Systems Volume II: Ambient Air Quality Monitoring Program." January 2017.

https://www3.epa.gov/ttnamti1/files/ambient/pm25/qa/Final%20Handbook%20Document%201_17.pdf.
Environmental Protection Agency. "Sulfur Dioxide Trends." Accessed April 11, 2019. https://www.epa.gov/air-trends/sulfur-dioxide-trends#sonat.
Gnoth, C, E Godehardt, P Frank-Herrmann, K Friol, Jurgen Tigges, and G Freundl. "Definition and Prevalence of Subfertility and Infertility." *Human Reproduction* 20(5). May 2005.
Gray, Simone C., Sharon E. Edwards, Bradley D. Schultz, and Marie Lynn Miranda. "Assessing the Impact of Race, Social Factors and Air Pollution on Birth Outcomes: A Population-Based Study." *Environmental Health* 13:4, 2013.
Guo, Z, RB Mosley, J McBrian, and RC Fortmann. "Fine Particulate Matter Emissions From Candles." 2000. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NRMRL&dirEntryId=63556.
Jarvis, Debbie J., Gary Adamkiewicz, Marie-Eve Heroux, Regula Rapp,and Frank J. Kelly. "Nitrogen Dioxide." *WHO Guidelines for Indoor Air Quality: Selected Pollutants*. 2010. https://www.ncbi.nlm.nih.gov/books/NBK138707/.
Laumbach R, Meng Q, Kipen H. "What can individuals do to reduce personal health risks from air pollution?" *J Thorac Dis*. 2015;7(1):96-107.
Legro, Richard S., Mark V. Sauer, Gilbert L. MOttla, Kevin S. Richter, Xian Li, William C. Dodson, and Duanping Liao. "Effect of Air Quality on Assisted Human Reproduction." *Human Reproduction* 25:5, 2010.
Leiser, Claire L., Heidi A. Hanson, Kara Sawyer, Jacob Steenblik, Ragheed Al-Dulaimi, Troy Madsen, Karen Gibbing, James M. Hotaling, Yetunde Oluseye Ibrahim, James A. VanDerslice, and Matthew Fuller. "Acute Effects of Air Pollutants on Spontaneous Pregnancy Loss: A Case-Crossover Study. *Fertility and Sterility* 111(2). 2019.
Liu, Yuewei, Yun Zhou, Jixuan Ma, Wei Bao, Jingjing Li, Ting Zhou, Xiuqing Cui, Zhe Peng, Hai Zhang, Min Feng, Yuan Yuan, Yuanqi Chen, Xiji Huang, Yonggang Li, Yonggang Duan, Tingming Shi, Lei Jin, Li Wu. "Inverse Association Between Ambient Sulfur Dioxide Exposure and Semen Quality in Wihan, China." *Environmental Science and Technology* 51(21). 2017.
Mohallem, Soraya Vecci, Debora Ja de Araujo Lobo, Celia Regina Pesquero, Joao Vicente Assuncao, Paulo Afonso de Andre, Paulo Hilario Saldiva, and Marisa Dolhnikov. "Decreased Fertility in Mice Exposed to Environmental Air Pollution in the City of Sao Paulo." *Environmental Research* 92:2, June 2005.
National Institute for Occupational Safety and Health. "Reproductive Health and the Workplace: Smoke and Byproducts of Burning." Accessed April 11, 2019. https://www.cdc.gov/niosh/topics/repro/smoke.html.
National Institute of Health. "ToxTown: Carbon Monoxide." Accessed April 11, 2019. https://toxtown.nlm.nih.gov/chemicals-and-contaminants/carbon-monoxide.
Strosnider, Heather, Caitlin Kennedy, Michele Monti, and Fuyuen Yip. "Rural and Urban Differences in Air Quality, 2008–2012, and Community Drinking Water Quality, 2010–2015 — United States." Centers for Disease Control and Prevention *Surveillance Summaries* 66:13, 2017.
Selevan, Sherry G., Libor Borkovec, Valerie L. Slott, Zdena Zudova, Jiri Rubes, Donald P. Evenson, and Sally D. Perreault. "Semen Quality and Reproductive Health of Young Czech Men Exposed to Seasonal Air Pollution." *Environmental Health Perspectives* 108:9, 2000.

Thurston SW, Ryan L, Christiani DC, Snow R, Carlson J, You L, Cui S, Ma G, Wang L, Huang Y, et al. Petrochemical exposure and menstrual disturbances. Am J Indust Med. 2000;38:555–64.

Tomei, Gianfranco, Manuela Ciarrocca, Bruna Rita Fortunato, Assunta Capozzella, Maria Valeria Rosati, Daniela Cerratti, Enrico Tomao, Vincenza Anzelmo, Carlo Monti, and Francesco Tomei. "Exposure to Traffic Pollutants and Effects on 17-B-Estradiol (E2) in Female Workers." *International Archives of Occupational and Environmental Health* 80:1, 2006.

United States Census Bureau. "Methodology For National Intercensal Estimates." Accessed April 11, 2019. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/intercensal/intercensal-nat-meth.pdf.

United States Environmental Protection Agency. "Managing Air Quality – Air Pollutant Types: Common Air Pollutants." Last Updated September 1, 2017. https://www.epa.gov/air-quality-management-process/managing-air-quality-air-pollutant-types#com.

Walker, Reed W. "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce." *Quarterly Journal of Economics*, 2013.

Watanabe, Nobue. "Decreased Number of Sperms and Sertoli Cells in Mature Rats Exposed To Diesel Exhaust as Fetuses." *Toxicology Letters* 155(1). January, 2005.

Zou, B, FB Zhan, Y Zeng. "Maternal Sulfur Dioxide Exposure and the Risk of Low Birth-weight Babies." *Journal of Hygeine Research* 40(5):638-642. September 2011.