

Air pollution and birth weight: Evidence from extremely polluted places*

Maxim Massenkoff
Naval Postgraduate School

September 2023

Abstract

I document a surprising resilience to air pollution in the world's most polluted places, at least as measured by birth outcomes. Average birth weight and the incidence of low birth weight in some of the most polluted cities in the world, and in highly polluted US counties in the 1970s, are essentially identical to the contemporary US. This is puzzling since quasi-experimental studies find large negative impacts of air pollution on fetal health. I discuss several possible explanations.

*Email: maxim.massenkoff@nps.edu. Jalaluddin Khan provided excellent research assistance. I thank to Gantuya Dorj for help understanding Ulaanbaatar birth data. For helpful feedback and comments, I thank Janet Currie, Stefano DellaVigna, Nathan Franz, Jacob Marcus, David Molitor, Joshua Rosenberg, Evan Rose, Aartik Sarma, Dean Spears, Alex Tabarrok, Christine Tsang, and Reed Walker. All errors are my own.

1 Introduction

The effects of air pollution are hard to measure because exposure is not random (e.g., [Jbaily et al., 2022](#)). Several studies on birth weight and pollution use quasi-experimental approaches to get around this issue. [Alexander and Schwandt \(2022\)](#) use purchases of high-pollution Volkswagens as an instrumental variable and find particulate matter has large negative impacts on birth weight. [Currie and Walker \(2011\)](#) find that the installation of automated toll booths increased birth weight in the surrounding areas, implying a large effect of nitrogen dioxide (NO₂) and other pollutants.¹

These results carry significant implications for public health. Moving from a 10th to a 90th percentile county in terms of PM_{2.5} would more than double exposure. The results in [Alexander and Schwandt \(2022\)](#) suggest that this change would increase the chance of low birth weight by 40% and decrease birth weight by 142 grams—about half the 300-gram decrease observed during the Dutch famine ([Stein and Susser, 1975](#)). The effects in [Currie and Walker \(2011\)](#) are even larger.

Testing these predictions using birth outcomes in the contemporary US would not be helpful: the US has a relatively narrow range of pollution, and it's easy to imagine that people in more polluted places are selected, so confounds could mask the impact. However, because of the fundamentally biological mechanism and large estimates, extremely polluted cities elsewhere in the world may complement the quasi-experimental approaches.

In this paper, I collect data on birth weights in the most polluted places in the world and on US counties that suffered high rates of pollution in the 1970s. In contrast to what might be expected from the modern causal estimates, I uncover three surprises: (1) Birth weights, measured using average birth weight and low birth weight incidence, are normal in some of the world's most polluted cities. For instance, babies born in Beijing, where particulate matter is 5-8 times as bad as in the United States, weigh about the same as American newborns. (2) Infants born in the most polluted US counties in the 1970s (e.g., Allegheny County, Pennsylvania), where pollution was twice as high as the most polluted counties today, also have normal birth weights. (3) Despite sizeable decreases in ambient PM_{2.5}, NO₂, and other contaminants ([EPA, 2022b,a](#)), birth weights in these previously highly-polluted US counties have been stable over time.

¹[Alexander and Schwandt \(2022, Figure 6\(a\)\)](#) provide a useful summary of estimates of the effect on infant health. The automated toll booths decreased pollution through their impact on idling cars. Only NO₂ was consistently measured, but the change likely decreased other pollutants as well.

I consider several explanations for why the birth weights in these cities are higher than expected given the causal estimates, and why birth weights haven't improved in the formerly polluted parts of the US. The main question is whether, despite the observed normal birth weights, pollution is still quite harmful and something either masks its effects or makes these contexts unsuitable for testing these predictions. I address the measurement of pollution and birth weight, culling effects, maternal selection, obstetric practices, differences in infant viability, and the timing of exposure. I find that none can explain the large gap between the predicted and observed birth outcomes.

Ignoring birth weight, the bulk of the evidence still suggests that particulate matter affects other important aspects of health ([Landrigan et al., 2018](#)), so addressing air pollution likely remains an urgent priority. However, potential fetal impacts are especially concerning given the long-run effects ([Almond and Currie, 2011](#); [Currie et al., 2014](#); [Black et al., 2007](#)). It is possible that birth weight is a complicated proxy for measuring the damage wrought by pollution and other health insults. Indeed, [Goldin and Margo \(1989\)](#) unearthed birth records from “poor houses” in 1800s Philadelphia, finding average birth weights similar to birth weights today. In light of the uncertainties outlined here, larger-scale randomized trials of air purification could provide key evidence on this important issue.

This paper contributes to a literature on the harms of pollution. Birth measures such as average birth weight and the incidence of low birth weight are widely used health measures ([Almond et al., 2005](#)) and constitute a significant share of the outcomes studied in causal evidence on pollution.² This might be because birth weight is more consistently measured ([Currie, 2013](#)). Another reason to be especially concerned about these outcomes is the potential for long-run impacts on adults ([Almond and Currie, 2011](#)). But research on pollution has also found effects on outcomes not studied here, such as infant mortality (e.g., [Chay and Greenstone, 2003](#); [Heft-Neal et al., 2020](#)), old-age mortality (e.g., [Deryugina et al., 2019](#)), cardiovascular disease (e.g., [Liang et al., 2020](#)), and asthma (e.g., [Alexander and Schwandt, 2022](#)), to give just a few examples.

This paper also contributes to a literature on the external validity of empirical findings (e.g.,

²For example, in its background research for numerous grants addressing air quality, [Open Philanthropy \(2022\)](#) highlights birth-related outcomes in particular:

The lack of reliable household PM2.5 concentration data makes it difficult to confidently discern health effects. The available evidence indicates that negative health outcomes of household air pollution in South Asia may include low birth weight, preterm birth, and other conditions that are correlated with an increased risk of infant death.

See [Almond et al. \(2005\)](#) for a careful examination of the health consequences of low birth weight.

Rosenzweig and Udry, 2020; List, 2020; Vivalt, 2020; DellaVigna and Linos, 2022) and combining evidence across studies (e.g., Meager, 2022). While the evidence assembled here is different from other observational and quasi-experimental approaches, policymakers must aggregate from disparate sources to evaluate policies (e.g., US EPA, 2021). And reasoning from extremes can be useful. For example, scientists suspecting a link between smoking and cancer applied tobacco tar to lab animals and documented the resulting tumors (Proctor, 2012). Both cases revolve around biological determinants of health—the causal effects of education, for example, might vary more widely across contexts. But for biological factors like pollution, tests like this may provide key evidence on the plausibility of the mechanism and effect sizes.

2 Background: Quasi-experimental evidence on pollution and infant health

Alexander and Schwandt (2022) use the rollout of “cheating” Volkswagen cars to study the impact of particulate matter on birth outcomes. The authors show evidence that, conditional on a set of controls, sales of the problematic cars are as-good-as-random, and these cars have a large impact on pollution as measured by EPA sensors. They report several key findings, but I focus on their two-stage least-squares (2SLS) in terms of PM2.5 particulate matter for ease of interpretation.³ They find that a 1 point (or ten percent) increase in PM2.5 decreases birth weight by 23 grams and increases the probability low birth weight ($< 2,500$ grams) by 0.44 percentage points (Table A.8).

3 Evidence from extremely polluted cities

In this section, I collect data on pollution and birth weight from the most polluted cities in the world. The relationship between pollution and birth outcomes within this collection of cities does not capture the causal effect of pollution. Rather, each individual city is seen as a test of the predictions from the causal estimates. Residents of these cities likely inhale much more particulate matter on average. I find that given the amount of pollution, birth weights are surprisingly normal in places like Beijing.

3.1 Extrapolation

The average PM2.5 was 85 in Beijing in 2013 (WHO, 2016b) compared to around 10 in the US (see Appendix C). How much lower should we expect their average birth weight to be? In what follows

³PM2.5 refers to particulate matter that is 2.5 micrometers in diameter or smaller. It is typically measured in micrograms per cubic meter ($\mu\text{g}/\text{m}^3$). For brevity, I leave out $\mu\text{g}/\text{m}^3$ and refer to it as PM2.5.

I use the relatively conservative estimates from [Alexander and Schwandt \(2022\)](#) to extrapolate. The effects measured in [Currie and Walker \(2011\)](#) are 80% larger (see [Alexander and Schwandt, 2022](#), Figure 6(A)), although it did not directly measure PM2.5.

Birth weight Based on the crudest linear interpretation of the 2SLS estimates in [Alexander and Schwandt \(2022\)](#)⁴, birth weight should be $(85 - 10) * 23 = 1,725$ grams or 52% lower in Beijing compared to the US. This is an unrealistically large effect. An obvious alternative is to recast the estimates in log terms, noting however that [Alexander and Schwandt \(2022\)](#) explicitly test for non-linear effects in the US context and find no evidence of this (Table 5).

Abstracting away from all factors besides pollution, the authors' estimates can be written as

$$BirthWeight_c = \alpha - 23 * PM25_c + e_c.$$

where $BirthWeight_c$ is average birth weight in city c , α is a constant, $PM2.5_c$ is the pollution in city c (or the US), and e_c is the error term. Since average PM 2.5 in the US is about 10, a 1-unit increase in PM 2.5 is a 10 percent increase. So birth weight goes down by 23 grams for every 10 percent increase in PM 2.5, or 2.3 grams for every 1 percent increase. This relationship could be captured in a linear-log specification as

$$BirthWeight_c = \alpha - 230 * \ln(PM25_c) + e_c.$$

The predicted birth weight in Beijing, relative to the US, is thus:

$$\Delta_{US-Beijing} = 230 * \ln(85/10) = 492 \text{ grams.} \quad (1)$$

This is a more plausible estimate, but is still large, corresponding to a roughly one standard deviation decrease. While there may not be a gold standard study of causal effects of fetal insults, [Goldin and Margo \(1989\)](#) note that a common point of comparison is the effects of the Dutch famine on birth weights, a decrease of about 300 grams ([Stein et al., 1975](#); [Stein and Susser, 1975](#)). In between-sibling designs, heavy smoking during pregnancy is associated with a 226 gram decrease in birth weight ([Juárez and Merlo, 2013](#)).⁵ So this conservative log extrapolation of the effects in [Alexander](#)

⁴These come from the IV estimates in Table A.8 Panel C column (1).

⁵[Currie et al. \(2009\)](#) perform the same design using data from New Jersey and get smaller estimates of the impact of

and Schwandt (2022) implies that living in Beijing is 60% worse than enduring a famine and twice as bad as smoking, at least based on the available estimates.

The authors also study incidence of low birth weight, defined as birth weight under 2,500 grams. They estimate that a 1-unit increase in PM2.5 causes a 0.44 percentage point increase in the probability of low birth weight. This means the difference in low birth weight between Beijing and the US should be $(85-10)*0.44 = 33$ percentage points. This linear extrapolation is similarly extreme. We can cast the pollution in log terms such that a 10% increase in PM2.5 (off a base of 10) leads to a 0.44 percentage point increase in low birth weight:

$$LBW_c = \delta + 0.044 * \ln(PM25_c) + e_c.$$

So the rate of low birth weight in Beijing should be

$$\Delta_{US-Beijing} = 0.044 * \ln(85/10) = 0.094$$

or 9 percentage points higher compared to the US. I use these linear and log extrapolations in the analysis that follows.

3.2 Data

I next present birth weight data on some of the most polluted cities in the world, collated from different medical sources that sample from the city itself. I choose cities rather than larger geographic areas as the level of analysis in order to have the sharpest possible measure of exposure to pollution. (Alexander and Schwandt (2022) use the US EPA's county-level sensors.) I restricted to cities in countries with a GDP per capita of at least \$2,000. I use this income threshold to make the groups slightly more comparable. The poorest countries in the world have more than ten times the rate of infant mortality compared to Western countries; birth outcomes are worse across all categories in these places, likely due to factors beyond just pollution.

I provide details on every source in Appendix C. Sources were found by searching for the city name and the words “average birth weight” or “birth weight.” These statistics are not available for

smoking, a reduction of 61 grams for women smoking 10 cigarettes a day. Almond et al. (2005) estimate a 200 gram effect using propensity score matching.

most cities. I first searched for the most polluted cities from each country, provided the city had a PM 2.5 of 30 or more, about two times higher than the 99th percentile county-year level in the United States. If I could not find data on that city, I searched for the next most-polluted city in the country and so on. The data needed to be representative of the population sampled and include at least 100 births. A major difficulty is that medical researchers often make restrictions that render the sample unusable (e.g., pre-term or full-term infants only, or only infants without major birth defects). Several of the present studies exclude multiple births, which tend to have lower birth weight. For simplicity, I exclude them from the US to not bias the comparison in favor of non-US places.

In some cases, the statistics used for the highly polluted cities are less reliable and less representative compared to US birth records. For example, the reference for Beijing ([Su et al., 2016](#)) excludes mothers over 43 (this is above the 99th percentile of US mother ages), and some data is based on household surveys rather than administrative hospital data (e.g., the source for Tunis, [Sassi et al., 2019](#)). The Ulaanbaatar study, [Barn et al. \(2018b\)](#), excluded mothers who smoke. However, many of the sources listed in [Appendix C](#) are from hospitals or hospital networks, and restricting to such studies would not change the results.

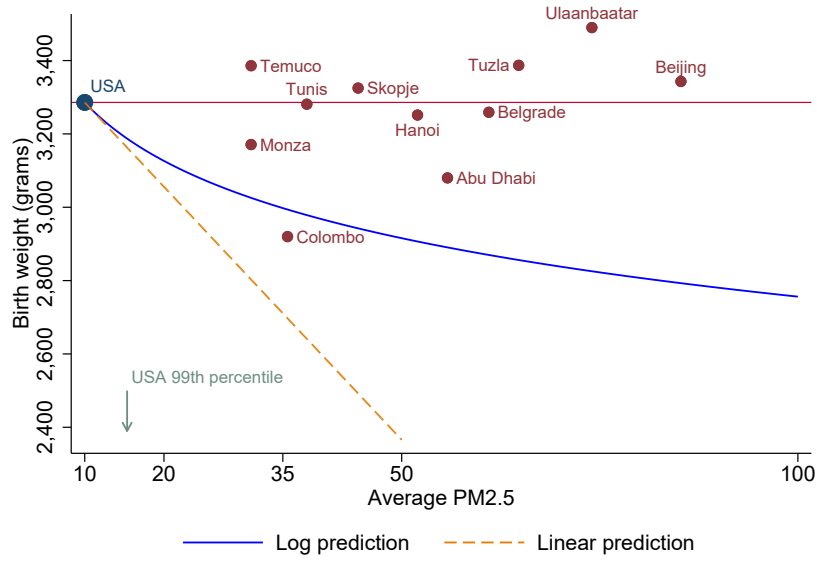
Pollution data comes from the World Health Organization ([WHO, 2016b](#)), and GDP data from the World Bank ([World Bank, 2022a](#)). The pollution data from WHO excludes some large countries, so in a few cases I used research papers for PM 2.5 estimates (see [Appendix C](#)). If the birth data was collected more than five years from the pollution data, I searched for a more proximate pollution estimate. I left out samples where I did not have a pollution estimate within five years.

3.3 Birth outcomes in highly polluted cities

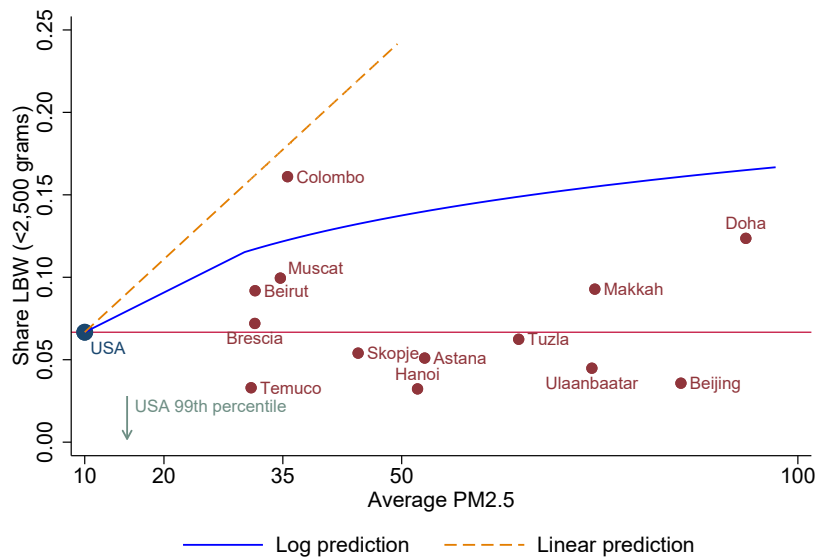
3.3.1 Average birth weight

I show the results for average birth weight in [Figure 1\(A\)](#) and [Table 1](#). To illustrate the procedure for one city: Ulaanbaatar has a reported annual mean PM2.5 of 74 ([WHO, 2016b](#)). This is almost five times higher than the 99th percentile county-year in the US. The linear prediction of its average birth weight, using the estimates in [Alexander and Schwandt \(2022\)](#), is 1,814 grams. Using the log prediction, its birth weight should be 2,826 grams. Its average birth weight is 3,343 grams, an estimate from a recent randomized trial of air purifiers and infant health ([Barn et al., 2018b](#)). This far exceeds either of the causal predictions.

Figure 1: Birth weights and pollution in the most polluted cities in the world



(A) Birth weight



(B) Low birth weight

Notes: This figure shows average birth weights for cities with extreme levels of pollution. In panel (A), the y-axis gives the average birth weight (in grams) in that city and the x-axis gives its level of pollution as measured by PM2.5. The solid blue line gives the log prediction derived from the causal estimates in [Alexander and Schwandt \(2022\)](#); the dashed gold line is the linear formulation (see Section 3.1). Panel (B) is analogous but with the share of low birth weight (< 2,500 grams) in the y-axis. Samples differ across the two panels due to limitations in which outcomes were available.

Table 1: Average birth weight in extremely polluted cities

City	Country	PM 2.5	Birth weight (grams)	Linear prediction	Log prediction
Beijing	China	85	3,343	1,555	2,793
Ulaanbaatar	Mongolia	74	3,490	1,814	2,826
Tuzla	Bosnia and Herzegovina	65	3,387	2,026	2,856
Belgrade	Serbia	61	3,259	2,113	2,870
Abu Dhabi	United Arab Emirates	56	3,080	2,233	2,891
Hanoi	Vietnam	52	3,251	2,320	2,907
Skopje	North Macedonia	45	3,325	2,492	2,942
Tunis	Tunisia	38	3,281	2,641	2,979
Colombo	Sri Lanka	36	2,920	2,697	2,994
Temuco	Chile	31	3,386	2,803	3,026
Monza	Italy	31	3,171	2,803	3,026
Extreme city average		52	3,263	2,310	2,921
Nationwide	USA	10	3,286	3,286	3,286

Notes: PM 2.5 data is from the World Health Organization ([WHO, 2016b](#)). Birth weight sources are described in [Appendix C](#). The math underlying the predictions is given in Section 3.1. The Extreme city average weights each city equally.

This is true of most of the considered cities: the average birth weight in the city substantially outstrips the predictions from the model, shown in the two columns on the right in [Table 1](#). To take another example, Beijing is predicted to have a birth weight of 2,793 grams. However, its average birth weight actually exceeds that of the US, at 3,343 grams. The equally-weighted average birth weight across all of these cities, shown at the bottom of [Table 1](#), is just slightly lower than in the United States.

In [Figure 1\(A\)](#), I give a visual representation of these findings. In most cases, the birth weights lie substantially above even the log-predicted birth weight. All cities except Colombo have higher average birth weights than the log prediction. The birth weights are also surprisingly large if we impose a floor on the predicted birth weight after PM2.5 passes 20. In other words, while there is surely measurement error in the pollution numbers, these birth weights would still exceed the causal predictions if these cities' true PM2.5 were only 20. This could also address the possibility that the effects of pollution are concave: that is, the effects of pollution decrease as pollution increases as suggested by [Miller et al. \(2021\)](#).

3.3.2 Low birth weight

Average birth weight may not be the appropriate measure if pollution makes extreme outcomes more likely while leaving most of the birth weight distribution unchanged. I collected data on the incidence

Table 2: Low birth weight incidence in extremely polluted cities

City	Country	PM 2.5	LBW	Linear prediction	Log prediction
Doha	Qatar	93	0.12	0.44	0.16
Beijing	China	85	0.04	0.40	0.16
Makkah	Saudi Arabia	74	0.09	0.35	0.15
Ulaanbaatar	Mongolia	74	0.04	0.35	0.15
Tuzla	Bosnia and Herzegovina	65	0.06	0.31	0.15
Astana	Kazakhstan	53	0.05	0.26	0.14
Hanoi	Vietnam	52	0.03	0.25	0.14
Skopje	North Macedonia	45	0.05	0.22	0.13
Colombo	Sri Lanka	36	0.16	0.18	0.12
Muscat	Oman	35	0.10	0.18	0.12
Beirut	Lebanon	32	0.09	0.16	0.12
Brescia	Italy	31	0.07	0.16	0.12
Temuco	Chile	31	0.03	0.16	0.12
Extreme city average		52	0.07	0.25	0.14
Nationwide	USA	10	0.07	0.07	0.07

Notes: PM 2.5 data is from the World Health Organization ([WHO, 2016b](#)) unless noted in [Appendix C](#). LBW is the fraction of births with low birth weight (< 2,500 grams). Birth outcome sources are described in [Appendix C](#). The math underlying the predictions is given in Section 3.1. The Extreme city average weights each city equally.

of low birth weight, defined as weighing under 2,500 grams, using the same approach. These estimates were easier to find because it's a more commonly measured outcome. The results are shown in [Table 2](#) and [Figure 1\(B\)](#). They echo the findings with average birth weight. Low birth weight in most of the highly polluted cities is below the rate for the US, including in Ulaanbaatar and Beijing.

To take one example, WHO reports that PM 2.5 in Brescia, Italy is 31. The log formulation of estimates from [Alexander and Schwandt \(2022\)](#) predicts a low birth weight incidence of 13% based on this level of pollution. My source for low birth weight incidence is a vital statistics report from Italy. It reports a low birth weight incidence of 7% among parents living in Brescia ([ATS Brescia, 2020](#)), which is about the same as the US.

Note that in this analysis, as mentioned above, I make the conservative adjustment of removing all multiple births from the US sample. In the US data, this decreases the low birth weight incidence from 8.3% to 6.7% ([Appendix C](#)). This is to deal with the fact that some of the international samples omit multiple births, and for others it is unclear (see [Appendix C](#)). Overall, this comparison is probably biased slightly toward finding higher rates of low birth weight in the international sample if most studies actually included multiple births.

3.4 Other international sources

In [Appendix D](#), I discuss other noteworthy sources which did not meet the inclusion criteria. In particular, the city of Teplice in the Czech Republic was once one of the most polluted cities in the world. In the early 1990s, the Czech public health department entered into a special collaboration with the US Environmental Protection Agency to study the health consequences of pollution there. These studies yielded detailed estimates of birth outcomes and exposure, and similarly show no clear harms as far as the incidence of low birth weight and pre-term birth ([Sram et al., 1996](#)). I also discuss [Rich et al. \(2015\)](#). It has better pollution and birth data than the Beijing source employed in the main analysis, and has similarly normal birth outcomes, but the authors excluded preterm births.

3.5 Preterm birth

[Alexander and Schwandt \(2022\)](#) and [Currie and Walker \(2011\)](#) also study preterm birth as an outcome, defined as a birth before 37 weeks of gestation. Some of the studies in the international sample also recorded this and I note the results in [Appendix C](#). In this case, there are fewer estimates, but the numbers give a similar impression: Beijing records a preterm birth rate of 5.0%, in Belgrade it is 8.7%, in Brescia it is 6.8%, and in Doha it is 14%. Except for Doha, these rates are lower than in the US. The preterm birth rate in 2021 was 10.5% ([CDC, 2023](#)) and about 8% for singletons ([Martin and Osterman, 2021](#)).

Together, these findings suggest that infants in highly-polluted areas around the world have surprising resilience to pollution—at least based on the most commonly used outcomes of birth weight and low birth weight. However, one potential risk is that these samples are somehow selected. For example, most studies are not population-level surveys and may consist of research hospitals serving richer communities. Next, I use US natality files to ask whether birth weights in previously highly polluted counties seem obviously impacted. This helps address selection concerns with the sparse international cities here and allows us to study other aspects of the birth such as gestation.

3.6 Historical birth weights in highly polluted US counties

The United States used to be much more polluted. [Chay and Greenstone \(2003, Figure 1\)](#) report that the national mean of total suspended particulates (TSP) was 93 micrograms per cubic meter in 1970

and just under 60 in 1990. Unfortunately, there's not a clean way to convert TSP to PM2.5. PM2.5 is restricted to smaller particles and the EPA only began tracking it in 1998 ([Voorheis et al., 2017](#)). [Lall et al. \(2004\)](#) use data from sensors with both measurements and find that the average ratio between PM2.5 and TSP particulates is 0.30. Using this crude conversion factor (cf. [Voorheis et al., 2017](#)), PM2.5 pollution in the US was around 28 in 1970 and 17 in 1990. I use this factor going forward noting that it is an approximation. Other pollutants such as sulfur dioxide, nitrogen oxide, and lead have all seen similarly dramatic decreases.

Geocoded birth weights from the 1970s to 2005 are available from the NBER's natality files ([NCHS, 2023](#)). These complement the data from non-US cities because there are minimal concerns about sample selection and it is easy to use the exact birth weights to measure the prevalence of low birth weight (< 2,500 grams) and very low birth weight (< 1,500 grams). Finally, information on the mothers allows us to probe the role of shifting maternal characteristics. Throughout I include twins from both the 1972 and contemporary samples, so these estimates are different from those in the international comparison.

In [Table 3](#), I show the county-level TSP along with the imputed PM2.5, the average birth weight, the percent of low birth weight and very low birth weight infants, and the number of births in the most 15 polluted counties from 1972. The results resemble those from the non-US cities: The birth weights are surprisingly similar to the 2019 US sample, despite the vast difference in exposure to particulate pollution. For example, Allegheny, PA recorded the highest pollution that year, with an imputed PM2.5 of 43. But its mean birth weight and low birth weight incidence, taken based on 9,378 infants born there in 1972, are all no worse than in the country-level aggregates for the US in 2019. Notably, the nitrogen dioxide exposure in Allegheny is also extreme: 64 parts per billion (ppb) compared to a median of under 10 ppb in the contemporary US ([US EPA, 2017](#), Figure 2.4)

The second to last row shows that the average birth weight in the polluted county sample is 3,286 grams, almost identical to the 2019 US average. We see a similar pattern in the two indicators for very small infants. The infants in 1972 are less likely to be low birth weight (7.1% in the 1972 sample vs. 8.3% in 2019) or very low birth weight (1.0% vs. 1.4%). Finally, all counties have lower incidence of preterm birth than 2019 data.

These estimates are at the county level, but the US natality data includes the city of residence for

Table 3: Pollution and birth outcomes in the most polluted US counties in 1972

Place	TSP	Est. PM 2.5	Birth weight	% LBW	% VLBW	% Preterm	N
Allegheny, PA	144	43	3,271	6.9	1.1		9,378
El Paso, TX	140	42	3,268	6.8	0.7		4,509
Riverside, CA	140	42	3,333	6.5	0.9	9.5	3,599
Webb, TX	137	41	3,327	5.8	1.1		1,140
Madison, IL	135	41	3,333	6.2	1.2	8.4	2,164
Tulare, CA	132	40	3,377	5.5	0.5	9.4	1,767
Kern, CA	133	40	3,302	6.9	0.9	8.2	3,006
San Bernardino, CA	132	39	3,316	7.0	1.1	8.7	5,580
Washington, PA	129	39	3,290	7.6	1.2		1,456
Mahoning, OH	130	39	3,287	7.7	1.1	8.6	2,286
Beaver, PA	127	38	3,257	8.5	1.4		1,399
Pinal, AZ	127	38	3,299	7.4	1.0	10.4	793
Denver, CO	125	38	3,134	9.4	1.1	9.5	4,262
Scott, IA	128	38	3,427	4.1	1.2	9.7	1,170
Black Hawk, IA	125	37	3,328	7.9	1.5	8.5	957
1972 Combined	135	40	3,286	7.1	1.0	9.0	43,466
2019 USA	33*	10	3,254	8.3	1.4	12.1	3,753,815

Notes: This table shows pollution and birth weights for extremely polluted counties in the US in 1972. TSP is total suspended particles as reported in the EPA Air Quality Trends report for 1972. The PM2.5 column is estimated by multiplying TSP by 0.3, except for the 2019 USA row, where TSP is estimated using the same factor. Birth weight is from the NHCS natality file for 1972 and (for the USA row) 2019. LBW is defined as birth weight under 2,500 grams, VLBW is birth weight under 1,500 grams, preterm is defined as gestation < 37 weeks. Gestation is inconsistently recorded so I drop counties with less than 50 percent coverage. N is the sample size in that place.

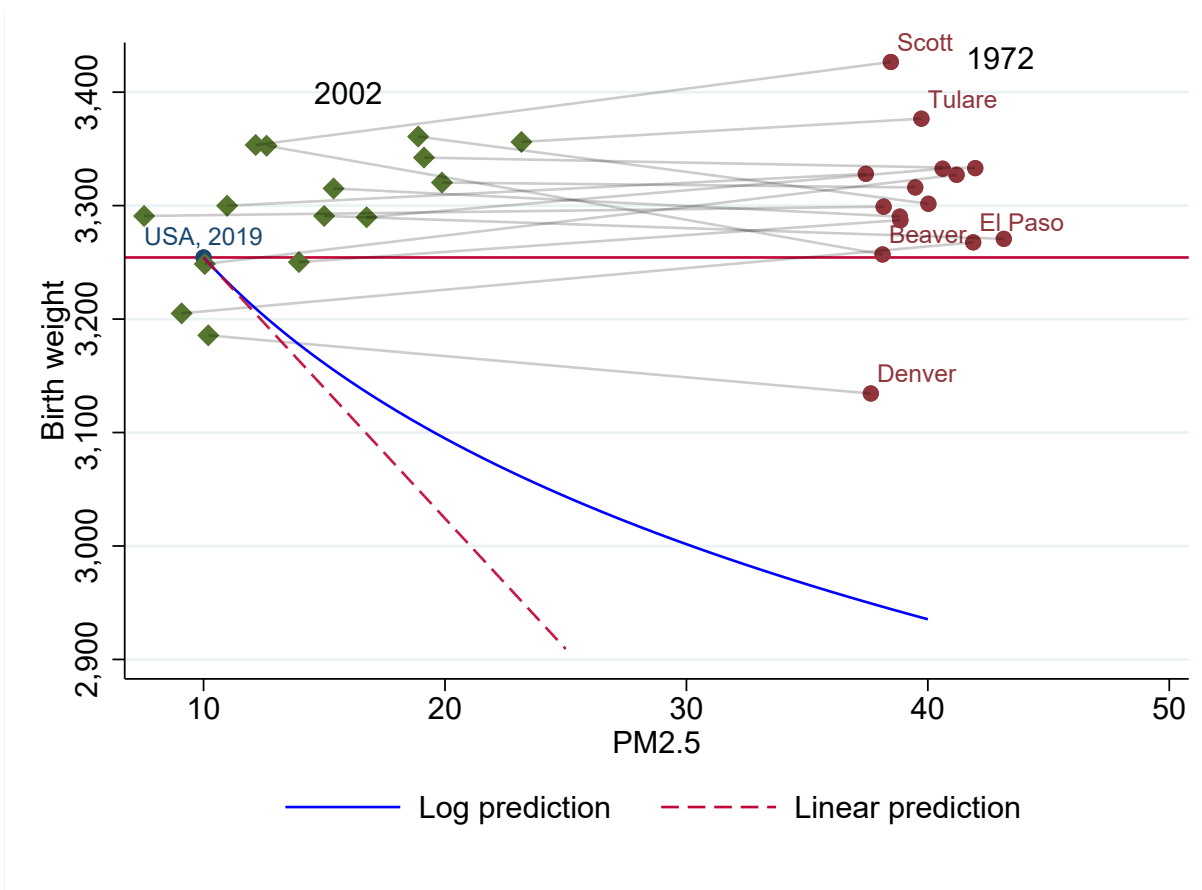
cities with populations exceeding 250,000 (NCHS, 1972). In theory, city-level pollution should be more representative of exposure because it is measured in a smaller area. Seven cities register as extremely polluted (TSP over 100) and are large enough to be indicated in the natality files: Birmingham, Denver, Detroit, El Paso, Newark, Phoenix, and Wichita. I show results for them in Figure A.1 with the estimates reported in Table B.1. These cities similarly surpass the predictions of the causal evidence although their birth outcomes are slightly worse, driven by Newark and Detroit. In particular, the low birth weight incidence in the combined group is 9.9% compared to 8.3% in the contemporary US. But the more conservative log formulation predicts that low birth weight incidence should be 14%.

I next use the county-level data to study how the change in pollution from 1972 to 2002 might have improved birth outcomes.⁶ Figure 2 shows two observations for each county in the high-pollution sample: one from 1972 and one from 2002. The maroon circles show data points from 1972, the same

⁶I stop at 2002 because the natality files switch their birth weight coding in 2004 (and from 2005 onward the geocodes are only available in the limited access data).

as in Table 3. The green diamonds show the same data from 2002. On average, these counties saw a 64% decrease in PM2.5, from an (imputed) average of 39 to a (measured) average of 14.

Figure 2: Birth weights and pollution predictions in the most polluted US counties in 1972-2002



Notes: This figure shows county-level PM 2.5 measurements and average birth weight for 1972 and 2002 for the 15 most polluted counties from 1972. The math underlying the predictions is given in Section 3.1.

The decrease in PM2.5 should have led to increases in birth weight. Instead, most move laterally in the plot: PM2.5 decreased without any meaningful improvement in birth outcomes. The average percent change in birth weights is about zero.

Older US sources Data on birth weights before the 1960s is scarce, but there are a few notable examples. Goldin and Margo (1989) study birth records from the Almshouse Hospital in Philadelphia in 1848-1837. The overall mean birth weight is 3,377 grams, although it's unclear how polluted Philadelphia was at this time. It's striking, however, how these mothers were selected: "The majority of the

women were destitute, abandoned, and without other housing; they were, in other words, the poorest of the poor” ([Goldin and Margo, 1989](#), p. 370).

4 Potential explanations

These findings are inconsistent with causal estimates from [Alexander and Schwandt \(2022\)](#) and [Currie and Walker \(2011\)](#). Here I review potential explanations.

Pollution measurement One issue is the measurement of exposure in the international sample. The city PM2.5 estimates come from ground-level sensors ([WHO, 2016b](#)), which may not be informative of experienced pollution for the typical person. In the last column of [Table B.2](#), I show estimates of country-level urban PM2.5 exposure that integrate satellite measurements and spatial population weights ([WHO \(2023\)](#); based on [Shaddick et al. \(2018\)](#)). These are not provided at the city level. But the exposure in urban areas in these countries generally confirm that pollution is far worse in these places. For example, [WHO \(2023\)](#) estimates that the PM2.5 exposure in urban places is 49.8 in China and 47.1 in Mongolia.

In [Section E.1](#), I review studies using personal sensors to measure pollution; these also suggest severe levels of exposure. If, despite these lines of evidence, people in the highly polluted samples actually experience moderate levels of exposure, it would have significant implications for how we quantify the damages of air pollution. These measurements undergird estimates of the global health burden of pollution (e.g., [Landrigan et al., 2018](#); [WHO, 2016a](#)). This issue appears less relevant in the historical US sample, where it seems widely acknowledged that the harms from pollution were far greater than today (e.g., [Chay and Greenstone, 2003](#)).

Birth weight measurement Birth weight may be inconsistently measured in middle-income countries, with issues like missing home births and heaping at multiples of 500 grams (see the review in e.g., [K’Oloo et al., 2023](#)). However, the highly polluted international sample includes hospital data from a modern study in Monza, Italy ([Ornaghi et al., 2022](#)), hospital data from an older study in Belgrade ([Maric et al., 2010](#)), a modern government health report from Brescia, Italy ([ATS Brescia, 2020](#)), a government report from Macedonia ([Stojanovska et al., 2006](#)), a health and demographic surveillance site set up by an international team in Hanoi ([Tran et al., 2012](#)), and an international trial of air puri-

fiers in Ulaanbaatar that tracked mothers from early in pregnancy ([Barn et al., 2018b](#)). I also noted the study of Teplice in [Appendix D](#) ([Sram et al., 1996](#)) and a detailed study of Beijing births which reports similar outcomes ([Rich et al., 2015](#)). It seems unlikely that across this array of higher-quality studies, issues in measurement led to a consistent underestimation of fetal harm.

Selection Another potential explanation is selection. Owing to the lower levels of economic development and worse health outcomes overall in the international sample, selection effects would seem to predict lower birth weights. I show in [Section F.1](#) that life expectancy is lower in most of these countries, and all rank lower on health and development. I discuss the demographics of the samples, where available, which also suggest negative selection relative to US births. In [Section F.2](#) I show that, within the US natality data, adjusting for maternal demographics does not change the results.

Culling effects Extreme pollution could induce a culling effect, disproportionately removing infants with worse birth outcomes from the population. In [Appendix G](#), I argue that culling effects, while difficult to assess, seem unlikely. First, culling does not seem to be happening during gestation or conception: The available data suggests that miscarriage rates are not much higher in polluted places. In particular, the Ulaanbaatar study ([Barn et al., 2018b](#)) explicitly tracked miscarriage rates, and these were around 2 percentage points higher than the US. Adding these births back into the Ulaanbaatar birth distribution with severely low birth weight would still leave Ulaanbaatar with higher average birth weight than the US. Another study of miscarriage in rural Kenya also shows relatively similar rates compared to the US ([Dellicour et al., 2016](#)).

Also, several papers from the international sample provide estimates of stillbirth rates: Monza (0.4%), Brescia (0.3%), Tuzla (0.7%) and Belgrade (1.3%) (see [Appendix C](#)). Except for Belgrade, these are lower than in the US (0.6%; [Gregory et al. \(2022\)](#)).

Culling could also happen at conception, but in vitro fertilisation success rates (a more controlled setting for studying the difficulty of conception) in Beijing resemble those in Europe. In general, such culling effects would contradict a vast literature which has never documented that pollution *increases* aggregate birth weight.

Obstetric practices Changing obstetric practices may have worked to decrease birth weight in the US ([Tilstra and Masters, 2020](#); [VanderWeele et al., 2012](#)). Also, the US and other developed countries use a lower threshold for the viability of very premature infants, which will lower its average birth weight relative to low- and middle-income countries. But in either case, adjusting for this using US data by calculating birth weight per week of gestation or removing extremely premature infants still leaves a large gap between observed and predicted birth outcomes ([Section G.3](#)).

Censoring the US birth distribution These potential explanations around miscarriage and viability can be viewed as a censoring of the birth weight distributions in the international samples relative to the US. While we do not have data on the full distributions in the non-US samples, we can use the contemporary US data to illustrate just how censored the samples would have to be for the predictions to line up with the causal estimates. I show in [Section G.4](#) that such censoring would need to be extreme—removing at least the bottom 5 percent of the birth weight distribution—to reconcile the estimates.

Timing of exposure The places studied here have or had stably high levels of pollution, while the causal literature focuses on shorter-term shocks. Could this explain the results? In [Appendix H](#), I discuss the evidence that short-term effects might be worse than long-term effects. If this is true, randomized controlled trials of air purifiers might be expected to show large benefits. But such trials have also failed to find major effects on birth outcomes, although these estimates are not powered to detect more moderate effects.

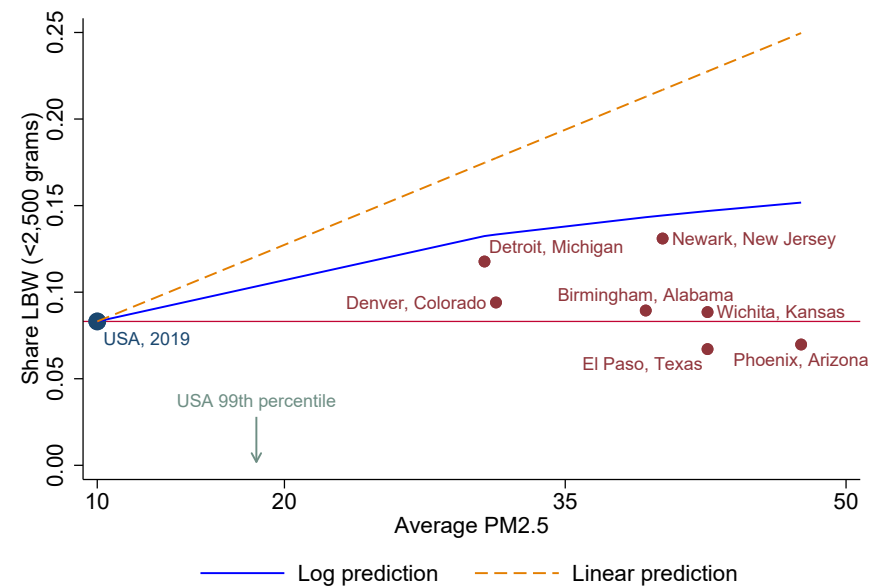
5 Discussion

This paper uses the causal effects of air pollution estimated in [Alexander and Schwandt \(2022\)](#) to predict birth outcomes in extremely polluted cities in middle-income countries and in US counties from the 1970s. I find that birth weights in highly polluted cities, which average five times the particulate matter of the contemporary United States, are surprisingly normal. So are birth outcomes in the most polluted US places in the 1970s, which, despite decreases in pollution, show no signs of improvement when I analyze outcomes in the same areas from 2002. The findings present a puzzle. Based on the causal estimates, infants in these extreme environments should have worse outcomes.

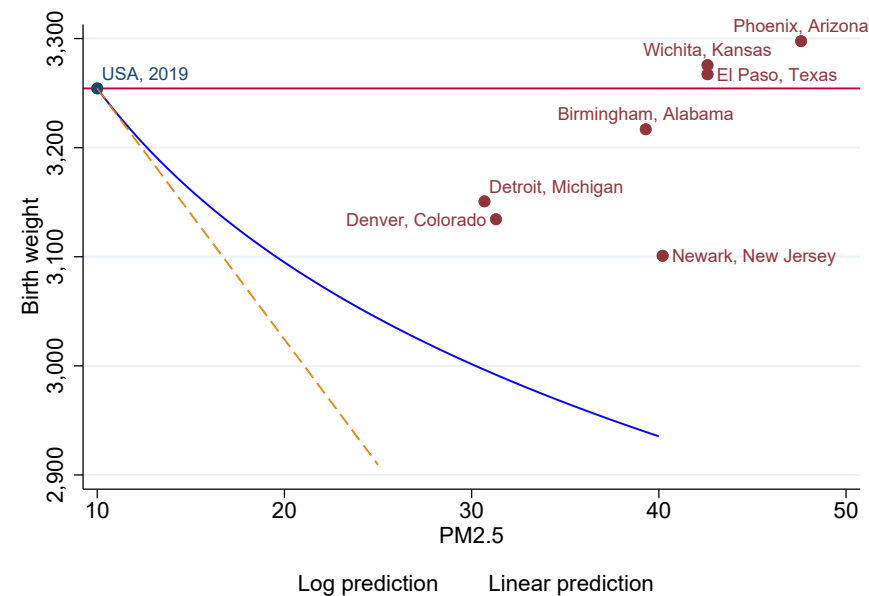
Pollution has many potential health effects ([Landrigan et al., 2018](#)), but impacts on fetal health are especially important because of the potential long-run consequences on life outcomes ([Almond and Currie, 2011](#)). Indeed, investment in child health has some of the best benefit-cost ratios compared to other policies ([Hendren and Sprung-Keyser, 2020](#)). Given the uncertainties outline here, a randomized study of air purification could provide key evidence on the causal effects of air pollution.

A Additional Figures

Figure A.1: City-level birth and pollution in the US, 1972-2002



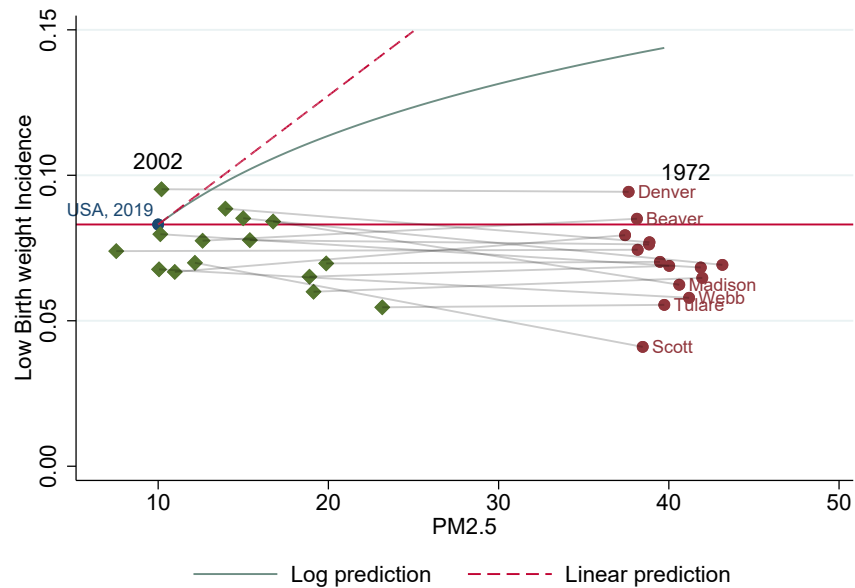
(A) Low birth weight incidence



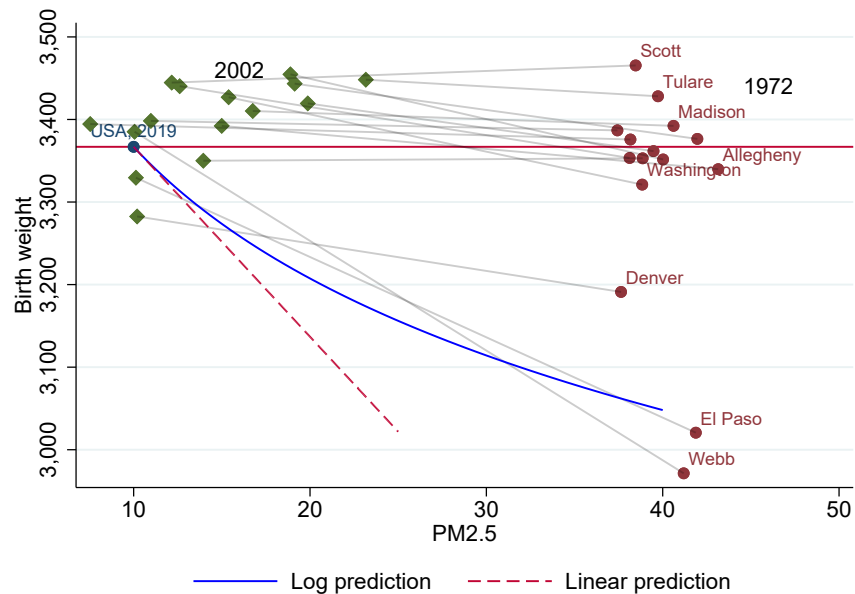
(B) Average birth weight

Notes: Birth outcomes from NCHS Natality Files, pollution data from EPA Air Quality Trends report. Low birth weight is defined as birth weight under 2,500 grams.

Figure A.2: Low birth weight and constant gestation birth weight, US counties 1972-2002



(A) Low birth weight

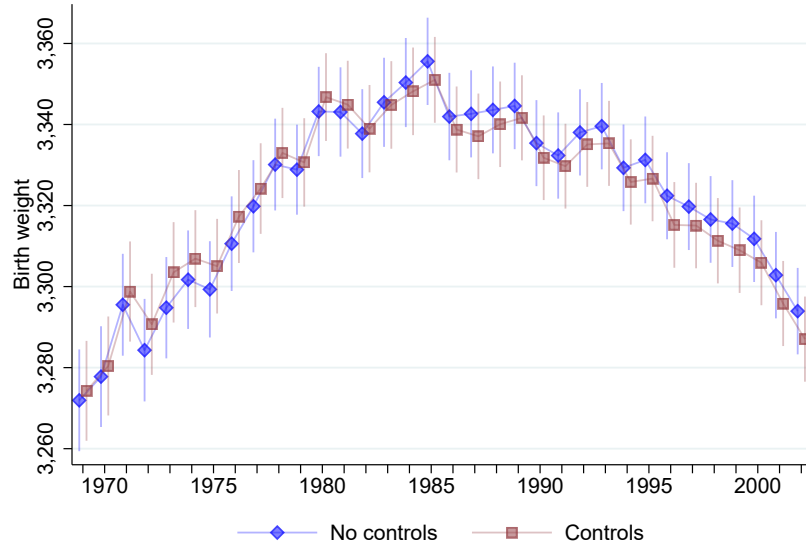


(B) Constant gestation birth weight

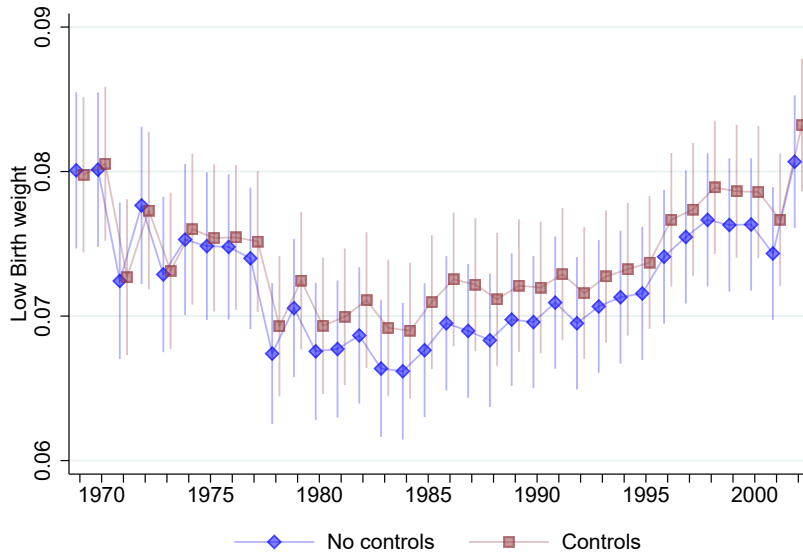
Notes: Birth outcomes from NCHS Natality Files, pollution data from EPA sensors. Low birth weight is defined as birth weight under 2,500 grams. The constant gestation birth weight is calculated at the infant level as:

$$40 * (\text{Birth Weight}) / (\text{Weeks of Gestation})$$

Figure A.3: Trends in US birth weights with and without controlling for maternal characteristics



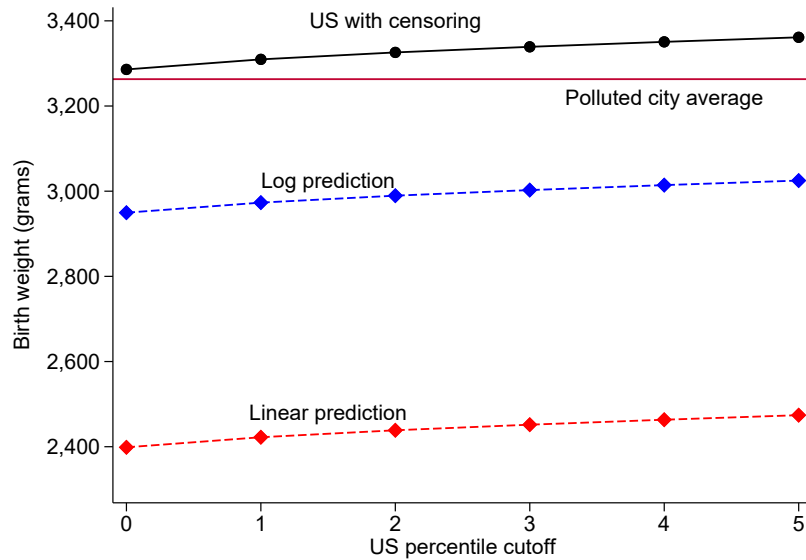
(A) Birth weight



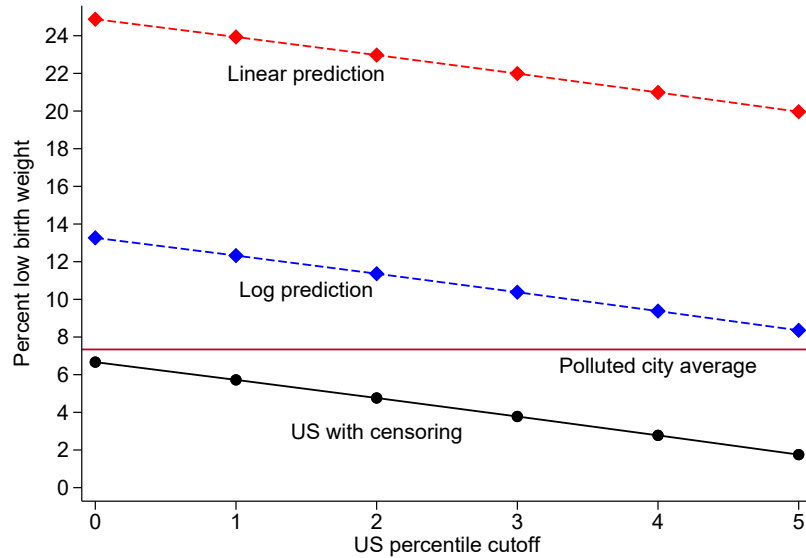
(B) Low birth weight

Notes: These panels show the estimated coefficients on year dummies in regressions using a 1 percent sample of the natality files from 1968 to today. The controlled regression includes fixed effects for mother age, race, live birth order, and state of residence. In Panel (A), the outcome is birth weight in grams. In Panel (B), the outcome is an indicator for being low birth weight (<2,500 grams). Source is the US Natality files ([NCHS, 2023](#)).

Figure A.4: Results censoring the US birth weight distribution



(A) Birth weight



(B) Low birth weight

Notes: These panels show how the international sample's birth weights would compare to their predicted values when censoring the bottom 0 to 5 percent of the US birth weight distribution. In Panel (A), the solid black line shows how average US birth weight would change removing the bottom X percent of its distribution, where X is the value in the x-axis. The flat red line shows the average birth weight across the sampled cities. The two prediction lines show how the predictions from the causal estimates would change when the censored US estimate is used as the reference. Panel (B) is analogous, but with percent low birth weight as the outcome.

B Additional Tables

Table B.1: Birth weights and pollution in the most polluted US cities, 1972

Place	TSP	PM 2.5	Birth weight	% LBW	% VLBW	N
Phoenix, Arizona	159	48	3,298	7.0	0.9	5,583
El Paso, Texas	142	43	3,267	6.7	0.7	4,112
Wichita, Kansas	142	43	3,276	8.8	1.0	2,407
Newark, New Jersey	134	40	3,101	13.1	2.1	3,879
Birmingham, Alabama	131	39	3,217	8.9	1.0	3,094
Detroit, Michigan	102	31	3,151	11.8	2.2	13,511
Denver, Colorado	104	31	3,134	9.4	1.1	4,262
1972 Combined	124	37	3,193	9.9	1.5	36,848
USA, 2019	33*	10	3,254	8.3	1.4	3,753,815

Notes: This table shows pollution and birth weights for extremely polluted counties in the US in 1972. TSP is total suspended particles as reported in the EPA Air Quality Trends report for 1972. The PM2.5 column is estimated by multiplying TSP by 0.3, except for the USA row. Birth weight is from the NHCS natality file for 1972 and (for the USA row) 2019. LBW is defined as birth weight under 2,500 grams, VLBW is birth weight under 1,500 grams. N is the sample size in that place.

Table B.2: International sample with country-level attributes

City	Country	Life expectancy	GDP per capita	Birth weight (g)	LBW	Urban PM2.5
-	USA	78.8	65,280	3,286	0.07	8.4
Doha	Qatar	80.2	62,276		0.12	62.4
Abu Dhabi	United Arab Emirates	78.0	42,701	3,080		43.0
Brescia	Italy	83.2	33,642		0.07	17.3
Monza	Italy	83.2	33,642	3,171		17.3
Makkah	Saudi Arabia	75.1	23,140		0.09	57.0
Muscat	Oman	77.9	17,701		0.10	35.8
Temuco	Chile	80.2	14,742	3,386	0.03	22.6
Beijing	China	76.9	10,144	3,343	0.04	49.8
Astana	Kazakhstan	73.2	9,813		0.05	30.9
Beirut	Lebanon	78.9	7,578		0.09	24.0
Belgrade	Serbia	75.7	7,417	3,259		26.3
Tuzla	Bosnia and Herzegovina	77.4	6,120	3,387	0.06	31.9
Skopje	North Macedonia	75.8	6,070	3,325	0.05	30.8
Ulaanbaatar	Mongolia	69.9	4,405	3,490	0.04	47.1
Colombo	Sri Lanka	77.0	3,852	2,920	0.16	23.9
Tunis	Tunisia	76.7	3,575	3,281		26.4
Hanoi	Vietnam	75.4	2,715	3,251	0.03	21.1

Note: Life expectancy and GDP per capita data are at the country level for 2019. Source for life expectancy is [World Bank \(2022b\)](#) and source for GDP is [World Bank \(2022a\)](#). The estimate of country-level urban PM2.5 is from [WHO \(2023\)](#). Sources for city-level birth outcomes are given in [Appendix C](#).

References

- Abdulkader, Zeyad Mohannad, Sajjad ur Rahman, and Nuha Nimeri**, "The incidence of low birth weight and intrauterine growth restriction in relationship to maternal ethnicity and gestational age at birth—A PEARL study analysis from the State of Qatar," *Qatar Medical Journal*, 2013, 2012 (2), 10.
- Abuelgasim, Abdelgadir and Ashraf Farahat**, "Investigations on PM 10, PM 2.5, and their ratio over the emirate of Abu Dhabi, United Arab Emirates," *Earth Systems and Environment*, 2020, 4, 763–775.
- Aimukhametova, Gulzhan, Talshyn Ukybasova, Zaituna Hamidullina, Karlygash Zhubanysheva, Md Harun-Or-Rashid, Yoshitoku Yoshida, Hideki Kasuya, and Junichi Sakamoto**, "The impact of maternal obesity on mother and neonatal health: study in a tertiary hospital of Astana, Kazakhstan," *Nagoya journal of medical science*, 2012, 74 (1-2), 83.
- Alexander, Diane and Hannes Schwandt**, "The impact of car pollution on infant and child health: Evidence from emissions cheating," *Review of Economic Studies*, 2022.
- Almond, Douglas**, "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of political Economy*, 2006, 114 (4), 672–712.
- **and Janet Currie**, "Killing me softly: The fetal origins hypothesis," *Journal of economic perspectives*, 2011, 25 (3), 153–72.
- **, Kenneth Y Chay, and David S Lee**, "The costs of low birth weight," *The Quarterly Journal of Economics*, 2005, 120 (3), 1031–1083.
- Amegah, Adeladza K, Reginald Quansah, and Jouni JK Jaakkola**, "Household air pollution from solid fuel use and risk of adverse pregnancy outcomes: a systematic review and meta-analysis of the empirical evidence," *PloS one*, 2014, 9 (12), e113920.
- ATS Brescia**, "IL PERCORSO NASCITA NELLA ATS DI BRESCIA," *Sistema Socio Sanitario, Regione Lombardia, ATS Brescia*, 2020. <https://web.archive.org/web/20221124155551/https://www.ats-brescia.it/documents/3432658/71035536/02+evento+nascita+ATS+brescia+2020.pdf/86bdbd64-7848-a5f6-cd55-140f42f8ffa0>.

- Avalos, Lyndsay Ammon, Claudia Galindo, and De-Kun Li**, “A systematic review to calculate background miscarriage rates using life table analysis,” *Birth Defects Research Part A: Clinical and Molecular Teratology*, 2012, 94 (6), 417–423.
- Baccarelli, Andrea A, Yinan Zheng, Xiao Zhang, Dou Chang, Lei Liu, Katherine Rose Wolf, Zhou Zhang, John P McCracken, Anaité Díaz, Pier Alberto Bertazzi et al.**, “Air pollution exposure and lung function in highly exposed subjects in Beijing, China: a repeated-measure study,” *Particle and fibre toxicology*, 2014, 11 (1), 1–10.
- Badarch, Jargalsaikhan, James Harding, Emma Dickinson-Craig, Colleen Azen, Hilary Ong, Samantha Hunter, Pia S Pannaraj, Brigitta Szepesi, Tegshjargal Sereenendorj, Sumiya Davaa et al.**, “Winter Air Pollution from Domestic Coal Fired Heating in Ulaanbaatar, Mongolia, Is Strongly Associated with a Major Seasonal Cyclic Decrease in Successful Fecundity,” *International Journal of Environmental Research and Public Health*, 2021, 18 (5), 2750.
- Barn, Prabjit, Enkhjargal Gombojav, Chimedsuren Ochir, Bayarkhuu Laagan, Bolor Beejin, Gerel Naidan, Buyantushig Boldbaatar, Jargalsaikhan Galsuren, Tsogtbaatar Byambaa, Craig Janes et al.**, “The effect of portable HEPA filter air cleaners on indoor PM2.5 concentrations and second hand tobacco smoke exposure among pregnant women in Ulaanbaatar, Mongolia: The UGAAR randomized controlled trial,” *Science of the total Environment*, 2018, 615, 1379–1389.
- , —, —, Buyantushig Boldbaatar, Bolor Beejin, Gerel Naidan, Jargalsaikhan Galsuren, Bayarkhuu Legtseg, Tsogtbaatar Byambaa, Jennifer A Hutcheon et al.**, “The effect of portable HEPA filter air cleaner use during pregnancy on fetal growth: the UGAAR randomized controlled trial,” *Environment international*, 2018, 121, 981–989. <https://www.sciencedirect.com/science/article/pii/S0160412018311413>.
- Bekkar, Bruce, Susan Pacheco, Rupa Basu, and Nathaniel DeNicola**, “Association of air pollution and heat exposure with preterm birth, low birth weight, and stillbirth in the US: a systematic review,” *JAMA network open*, 2020, 3 (6), e208243–e208243.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes**, “From the cradle to the labor market? The effect of birth weight on adult outcomes,” *The Quarterly Journal of Economics*, 2007, 122 (1), 409–439.

- Blencowe, Hannah, Julia Krusevec, Mercedes De Onis, Robert E Black, Xiaoyi An, Gretchen A Stevens, Elaine Borghi, Chika Hayashi, Diana Estevez, Luca Cegolon et al.,** “National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis,” *The Lancet Global Health*, 2019, 7 (7), e849–e860.
- Borgini, Alessandro, Cristian Ricci, Martina Bertoldi, Paolo Crosignani, Andrea Tittarelli et al.,** “The EuroLifeNet study: how different microenvironments influence personal exposure to PM 2.5 among high-school students in Milan,” *Open Journal of Air Pollution*, 2015, 4 (01), 16.
- Bozzoli, Carlos, Angus Deaton, and Climent Quintana-Domeque,** “Child mortality, income and adult height,” 2007.
- Brown, Ryan and Duncan Thomas,** *On the long term effects of the 1918 US influenza pandemic*, BREAD, the Bureau for Research and Economic Analysis of Development, 2019.
- Bruckner, Tim and Ralph Catalano,** “The sex ratio and age-specific male mortality: evidence for culling in utero,” *American Journal of Human Biology: The Official Journal of the Human Biology Association*, 2007, 19 (6), 763–773.
- Catalano, Ralph, Tim Bruckner, Amy R Marks, and Brenda Eskenazi,** “Exogenous shocks to the human sex ratio: the case of September 11, 2001 in New York City,” *Human reproduction*, 2006, 21 (12), 3127–3131.
- CDC,** “Preterm birth,” *Division of Reproductive Health, National Center for Chronic Disease Prevention and Health Promotion*, 2023. <https://web.archive.org/web/20230611215203/https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm#>.
- Chamseddine, A, I Alameddine, M Hatzopoulou, and M El-Fadel,** “Seasonal variation of air quality in hospitals with indoor–outdoor correlations,” *Building and Environment*, 2019, 148, 689–700.
- Chay, Kenneth Y and Michael Greenstone,** “The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession,” *The quarterly journal of economics*, 2003, 118 (3), 1121–1167.

- Checkley, William, Kendra N Williams, Josiah L Kephart, Magdalena Fandiño-Del-Rio, N Kyle Steenland, Gustavo F Gonzales, Luke P Naeher, Steven A Harvey, Lawrence H Moulton, Victor G Davila-Roman et al.**, “Effects of a household air pollution intervention with liquefied petroleum gas on cardiopulmonary outcomes in Peru. A randomized controlled trial,” *American journal of respiratory and critical care medicine*, 2021, 203 (11), 1386–1397.
- Choe, SA, YB Jun, WS Lee, TK Yoon, and SY Kim**, “Association between ambient air pollution and pregnancy rate in women who underwent IVF,” *Human reproduction*, 2018, 33 (6), 1071–1078.
- Collivignarelli, Maria Cristina, Alessandro Abbà, Francesca Maria Caccamo, Giorgio Bertanza, Roberta Pedrazzani, Marco Baldi, Paola Ricciardi, and Marco Carnevale Miino**, “Can particulate matter be identified as the primary cause of the rapid spread of CoViD-19 in some areas of Northern Italy?,” *Environmental Science and Pollution Research*, 2021, 28, 33120–33132.
- Currie, Janet**, “Pollution and infant health,” *Child development perspectives*, 2013, 7 (4), 237–242.
- **and Hannes Schwandt**, “The 9/11 dust cloud and pregnancy outcomes: A reconsideration,” *Journal of Human Resources*, 2016, 51 (4), 805–831.
- **and Reed Walker**, “Traffic congestion and infant health: Evidence from E-ZPass,” *American Economic Journal: Applied Economics*, 2011, 3 (1), 65–90.
- **, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell**, “What do we know about short-and long-term effects of early-life exposure to pollution?,” *Annu. Rev. Resour. Econ.*, 2014, 6 (1), 217–247.
- **, Matthew Neidell, and Johannes F Schmieder**, “Air pollution and infant health: Lessons from New Jersey,” *Journal of health economics*, 2009, 28 (3), 688–703.
- Dejmek, Jan, Sherry G Selevan, Ivan Benes, Ivo Solanský, and Radim J Srám**, “Fetal growth and maternal exposure to particulate matter during pregnancy,” *Environmental health perspectives*, 1999, 107 (6), 475–480.
- DellaVigna, Stefano and Elizabeth Linos**, “RCTs to scale: Comprehensive evidence from two nudge units,” *Econometrica*, 2022, 90 (1), 81–116.

- Dellicour, Stephanie, George Aol, Peter Ouma, Nicole Yan, Godfrey Bigogo, Mary J Hamel, Deron C Burton, Martina Oneko, Robert F Breiman, Laurence Slutsker et al.**, “Weekly miscarriage rates in a community-based prospective cohort study in rural western Kenya,” *BMJ open*, 2016, 6 (4), e011088.
- Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif**, “The mortality and medical costs of air pollution: Evidence from changes in wind direction,” *American Economic Review*, 2019, 109 (12), 4178–4219.
- Dida, Gabriel O, Patrick O Lutta, Paul O Abuom, Tomislav Mestrovic, and Douglas N Anyona**, “Factors predisposing women and children to indoor air pollution in rural villages, Western Kenya,” *Archives of Public Health*, 2022, 80 (1), 1–13.
- Dong, Wei, Shan Liu, Mengtian Chu, Bin Zhao, Di Yang, Chen Chen, Mark R Miller, Miranda Loh, Junhui Xu, Rui Chi et al.**, “Different cardiorespiratory effects of indoor air pollution intervention with ionization air purifier: findings from a randomized, double-blind crossover study among school children in Beijing,” *Environmental Pollution*, 2019, 254, 113054.
- Donkelaar, Aaron Van, Randall V Martin, Chi Li, and Richard T Burnett**, “Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors,” *Environmental science & technology*, 2019, 53 (5), 2595–2611.
- Dorj, Gantuya, Angarmurun Dayan, Khuderchuluun Nanjid, Undram Lkhagvaa, and Chimed-suren Ochir**, “Is Air Pollution a Risk Factor for Low Birth Weight in Ulaanbaatar, Mongolia?,” *OALib*, 01 2014, 01, 1–6. https://www.scirp.org/pdf/OALibJ_2016032213273374.pdf.
- Dorji, Phurb**, “Average birth weight of term newborn babies: a hospital based study in Thimphu, Bhutan,” *Bhutan Health Journal*, 2015, 1 (1), 11–14.
- Drake, Patrick, Anne K. Driscoll, and T.J. Mathews**, “Cigarette Smoking During Pregnancy: United States, 2016,” NCHS Data Brief 305, National Center for Health Statistics, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services February 2018.

- Eilerts-Spinelli, Hallie, Julio Romero Prieto, Julie Ambia, Sammy Khagayi, Chodziwadziwa Kabudula, Jeffrey W Eaton, and Georges Reniers, "Evaluating pregnancy reporting in Siaya Health and Demographic Surveillance System through record linkage with ANC clinics," *International Journal of Population Data Science*, 2022, 7 (4).
- EPA, "Nitrogen Dioxide Trends (August 1, 2022)," *Air trends*, 2022. <https://web.archive.org/web/20221109003059/https://www.epa.gov/air-trends/nitrogen-dioxide-trends>.
- , "Particulate Matter (PM2.5) Trends (August 1, 2022)," *Air trends*, 2022. <https://www.epa.gov/air-trends/particulate-matter-pm25-trends>.
- Floris, Joël, Laurent Kaiser, Harald Mayr, Kaspar Staub, and Ulrich Woitek, "Investigating survivorship bias: the case of the 1918 flu pandemic," *Applied Economics Letters*, 2021, pp. 1–6.
- Ganbat, Gantuya, Tseren-Ochir Soyol-Erdene, Batbayar Jadamba et al., "Recent improvement in particulate matter (PM) pollution in Ulaanbaatar, Mongolia," *Aerosol and Air Quality Research*, 2020, 20 (10), 2280–2288.
- General Authority for Statistics, "2016 Demography Survey," *Population and vitality surveys*, 2016. https://web.archive.org/web/20220419081906/https://www.stats.gov.sa/sites/default/files/en-demographic-research-2016_2.pdf.
- GiveWell, "Clean Cookstoves," *Research on programs - Interim Intervention Report*, 2022. <https://www.givewell.org/international/technical/programs/clean-cookstoves>.
- Glass, Hannah C, Andrew T Costarino, Stephen A Stayer, Claire Brett, Franklyn Cladis, and Peter J Davis, "Outcomes for extremely premature infants," *Anesthesia and analgesia*, 2015, 120 (6), 1337.
- Goldin, Claudia and Robert A Margo, "The poor at birth: Birth weights and infant mortality at Philadelphia's Almshouse Hospital, 1848–1873," *Explorations in Economic History*, 1989, 26 (3), 360–379.
- Gregory, Elizabeth C.W., Claudia P. Valenzuela, and Donna L. Hoyert, "Fetal Mortality: United States, 2020," *National Vital Statistics Reports*, 2022, 71 (4). M.P.H., Division of Vital Statistics.

- Grippo, Alexandra, Jun Zhang, Li Chu, Yanjun Guo, Lihua Qiao, Ajay A Myneni, and Lina Mu,** "Air pollution exposure during pregnancy and spontaneous abortion and stillbirth," *Reviews on environmental health*, 2018, 33 (3), 247–264.
- Halilović, Jasminka and Amela Begić,** "General Characteristics of Newborns from the Area of Tuzla Canton, Born in the Course of the 2007," *Collegium antropologicum*, 2015, 39 (2), 345–351.
- Han, Tao, Dan Wang, Wenyu Xie, Changgen Liu, Qian Zhang, Zhichun Feng, and Qiuping Li,** "Obstetricians' Attitudes Toward the Treatment of Extremely Preterm Infants in China," *JAMA Network Open*, 2022, 5 (9), e2233511–e2233511.
- Hayden, Dean, Maria Esterlita Villanueva-Uy, Maria Katrina Mendoza, and Dominic Wilkinson,** "Resuscitation of preterm infants in the Philippines: a national survey of resources and practice," *Archives of Disease in Childhood-Fetal and Neonatal Edition*, 2020, 105 (2), 209–214.
- He, Dan, Xuying Zhang, Ya'er Zhuang, Zhili Wang, and Yu Jiang,** "China fertility report, 2006–2016," *China population and development studies*, 2019, 2 (4), 430–439.
- Heft-Neal, Sam, Jennifer Burney, Eran Bendavid, Kara K Voss, and Marshall Burke,** "Dust pollution from the Sahara and African infant mortality," *Nature Sustainability*, 2020, 3 (10), 863–871.
- Hendren, Nathaniel and Ben Sprung-Keyser,** "A unified welfare analysis of government policies," *The Quarterly Journal of Economics*, 2020, 135 (3), 1209–1318.
- Hesselmar, B, N Aberg, B Aberg, B Eriksson, and B Björkstén,** "Does early exposure to cat or dog protect against later allergy development?," *Clinical and experimental allergy: journal of the British Society for Allergy and Clinical Immunology*, 1999, 29 (5), 611–617.
- Hug, Lucia, Danzhen You, Hannah Blencowe, Anu Mishra, Zhengfan Wang, Miranda J Fix, Jon Wakefield, Allisyn C Moran, Victor Gaigbe-Togbe, Emi Suzuki et al.,** "Global, regional, and national estimates and trends in stillbirths from 2000 to 2019: a systematic assessment," *The Lancet*, 2021, 398 (10302), 772–785.
- Jack, Darby W, Kenneth Ayuurebobi, Carlos F Gould, Ellen Boamah-Kaali, Alison G Lee, Mohammed Nuhu Mujtaba, Steven Chillrud, Seyram Kaali, Ashlinn K Quinn, Stephaney Gyaase**

- et al., “A cluster randomised trial of cookstove interventions to improve infant health in Ghana,” *BMJ global health*, 2021, 6 (8), e005599.
- Janssen, Patricia A, Paul Thiessen, Michael C Klein, Michael F Whitfield, Ying C MacNab, and Sue C Cullis-Kuhl**, “Standards for the measurement of birth weight, length and head circumference at term in neonates of European, Chinese and South Asian ancestry,” *Open Medicine*, 2007, 1 (2), e74.
- Jayatissa, Renuka and SMM Hossaine**, “Nutrition and food security assessment in Sri Lanka 2009,” *Medical Research Institute, Sri Lanka*, 2010. <https://documents.wfp.org/stellent/groups/public/documents/ena/wfp245828.pdf>.
- Jbaily, Abdulrahman, Xiaodan Zhou, Jie Liu, Ting-Hwan Lee, Leila Kamareddine, Stéphane Verguet, and Francesca Dominici**, “Air pollution exposure disparities across US population and income groups,” *Nature*, 2022, 601 (7892), 228–233.
- Jorquera, Héctor, Francisco Barraza, Johanna Heyer, Gonzalo Valdivia, Luis N Schiappacasse, and Lupita D Montoya**, “Indoor PM_{2.5} in an urban zone with heavy wood smoke pollution: The case of Temuco, Chile,” *Environmental Pollution*, 2018, 236, 477–487.
- Juárez, Sol Pía and Juan Merlo**, “Revisiting the effect of maternal smoking during pregnancy on offspring birthweight: a quasi-experimental sibling analysis in Sweden,” *PloS one*, 2013, 8 (4), e61734.
- Katz, Joanne, James M Tielsch, Subarna K Khattri, Laxman Shrestha, Patrick Breysse, Scott L Zeger, Naoko Kozuki, William Checkley, Steven C LeClerq, and Luke C Mullany**, “Impact of improved biomass and liquid petroleum gas stoves on birth outcomes in rural Nepal: results of 2 randomized trials,” *Global Health: Science and Practice*, 2020, 8 (3), 372–382.
- Kenessariyev, Ussen, Alexander Golub, Michael Brody, Askhat Dosmukhametov, Meiram Amrin, Aya Erzhanova, Dinara Kenessary et al.**, “Human health cost of air pollution in Kazakhstan,” *Journal of Environmental Protection*, 2013, 4 (08), 869.
- Kenya Ministry of Energy**, “Kenya Household Cooking Sector Study: Assessment of the Supply and Demand of Cooking Solutions at the Household Level,” *Online Resource*, 2019. <https://eedadvisory.com/wp-content/uploads/2020/09/MoE-2019-Kenya->

[Cooking-Sector-Study-compressed.pdf](https://web.archive.org/web/20210630082629/https://eedadvisory.com/wp-content/uploads/2020/09/MoE-2019-Kenya-Cooking-Sector-Study-compressed.pdf) (accessed on 5/28/2023) Archived: <https://web.archive.org/web/20210630082629/https://eedadvisory.com/wp-content/uploads/2020/09/MoE-2019-Kenya-Cooking-Sector-Study-compressed.pdf>.

Kosirinond, Chanyut and Gwi-Yeop Son, “Report on Thailand Gender-Disaggregated Statistics,” *Office of Women’s Affairs and Family Development Ministry of Social Development and Human Security*, 2008. <https://web.archive.org/web/20230624162812/https://www.undp.org/sites/g/files/zskgke326/files/migration/th/A-Gender-report-English.pdf>.

K’Oloo, Alloys, Evance Godfrey, Annariina M Koivu, Hellen C Barsosio, Karim Manji, Veneranda Ndesangia, Fredrick Omiti, Mohamed Bakari Khery, Everlyne D Ondieki, Simon Kariuki et al., “Improving birth weight measurement and recording practices in Kenya and Tanzania: a prospective intervention study with historical controls,” *Population health metrics*, 2023, 21 (1), 1–9.

Lall, Ramona, Michaela Kendall, Kazuhiko Ito, and George D Thurston, “Estimation of historical annual PM_{2.5} exposures for health effects assessment,” *Atmospheric Environment*, 2004, 38 (31), 5217–5226.

Landrigan, Philip J, Richard Fuller, Nereus JR Acosta, Olusoji Adeyi, Robert Arnold, Abdoulaye Bibi Baldé, Roberto Bertollini, Stephan Bose-O’Reilly, Jo Ivey Boufford, Patrick N Breyse et al., “The Lancet Commission on pollution and health,” *The lancet*, 2018, 391 (10119), 462–512.

Leung, Dennis YC, “Outdoor-indoor air pollution in urban environment: challenges and opportunity,” *Frontiers in Environmental Science*, 2015, 2, 69.

Li, Xiangyu, Shuqiong Huang, Anqi Jiao, Xuhao Yang, Junfeng Yun, Yuxin Wang, Xiaowei Xue, Yuanyuan Chu, Feifei Liu, Yisi Liu et al., “Association between ambient fine particulate matter and preterm birth or term low birth weight: an updated systematic review and meta-analysis,” *Environmental Pollution*, 2017, 227, 596–605.

Liang, Fengchao, Fangchao Liu, Keyong Huang, Xueli Yang, Jianxin Li, Qingyang Xiao, Jichun Chen, Xiaoqing Liu, Jie Cao, Chong Shen et al., “Long-term exposure to fine particulate matter

and cardiovascular disease in China," *Journal of the American College of Cardiology*, 2020, 75 (7), 707–717.

Liang, Lu, Peng Gong, Na Cong, Zhichao Li, Yu Zhao, and Ying Chen, "Assessment of personal exposure to particulate air pollution: the first result of City Health Outlook (CHO) project," *BMC public health*, 2019, 19 (1), 1–12.

Lippmann, Morton, Mark Frampton, Joel Schwartz, Douglas Dockery, Richard Schlesinger, Petros Koutrakis, John Froines, Andre Nel, Jack Finkelstein, John Godleski et al., "The US Environmental Protection Agency Particulate Matter Health Effects Research Centers Program: a midcourse report of status, progress, and plans.," *Environmental Health Perspectives*, 2003, 111 (8), 1074–1092.

List, John A, "Non est disputandum de generalizability? a glimpse into the external validity trial," Technical Report, National Bureau of Economic Research 2020.

Lucila, Rey-Ares, Irazola Vilma Edit, Althabe Fernando, Sobrino Edgardo, Mazzoni Agustina, Serón Pamela, Lanas Fernando, Calandrelli Matías, and Rubinstein Adolfo, "Lower respiratory tract infection in children younger than five years of age and adverse pregnancy outcomes related to household air pollution in Bariloche (Argentina) and Temuco (Chile)," *Indoor air*, 2016, 26 (6), 964.

Luong, Ly MT, Dung Phung, Peter D Sly, Lidia Morawska, and Phong K Thai, "The association between particulate air pollution and respiratory admissions among young children in Hanoi, Vietnam," *Science of the total environment*, 2017, 578, 249–255.

Maric, Nadja P, Bojana Dunjic, Dragan J Stojiljkovic, Dubravka Britvic, and Miroslava Jasovic-Gasic, "Prenatal stress during the 1999 bombing associated with lower birth weight—a study of 3,815 births from Belgrade," *Archives of women's mental health*, 2010, 13 (1), 83–89.

Martin, Joyce A and Michelle JK Osterman, "Exploring the Decline in the Singleton Preterm Birth Rate in the United States, 2019–2020," *NCHS Data Brief No. 430*, 2021. <https://www.cdc.gov/nchs/data/databriefs/db430.pdf>.

Mathews, Thomas J and Brady E Hamilton, "Mean age of mothers is on the rise: United States, 2000-2014.," *NCHS data brief*, 2016, (232), 1–8.

- Mathews, TJ, Brady E Hamilton et al.**, “Trend analysis of the sex ratio at birth in the United States,” *National vital statistics reports*, 2005, 53 (20), 1–17.
- Meager, Rachael**, “Aggregating distributional treatment effects: A Bayesian hierarchical analysis of the microcredit literature,” *American Economic Review*, 2022, 112 (6), 1818–47.
- Miller, Nolan, David Molitor, and Eric Zou**, “A causal concentration-response function for air pollution: evidence from wildfire smoke,” *NBER EEE Conference Paper*, 2021. https://conference.nber.org/conf_papers/f161570.pdf.
- Ministry of Health Sultanate of Oman**, “Annual Health Report 2011 A.D. 1432-1433 A.H.,” *Department of Health Information and Statistics, Directorate General of Planning, Ministry of Health, Sultanate of Oman*, 2012. Issued in 2012 (1121H/2341-2344AH).
- Moschandreas, Demetrios J, DJ Pelton, and J Zabransky**, *Comparison of indoor and outdoor air quality*, Electric Power Research Institute, 1981.
- Mukherjee, Sudeshna, Digna R Velez Edwards, Donna D Baird, David A Savitz, and Katherine E Hartmann**, “Risk of miscarriage among black women and white women in a US prospective cohort study,” *American journal of epidemiology*, 2013, 177 (11), 1271–1278.
- NCHS**, “NATALITY DETAIL TAPE DOCUMENTATION,” 1972-1977 *NATALITY FILES*, 1972. <https://data.nber.org/natality/1972/natl1972.pdf>.
- NCHS**, “Vital Statistics Natality Birth Data,” 2023. Natality Data from the National Vital Statistics System of the National Center for Health Statistics provided by NBER, accessed at <https://www.nber.org/research/data/vital-statistics-natality-birth-data>.
- Nguyen, Huong Thu, Bo Eriksson, Liem Thanh Nguyen, Chuc Thi Kim Nguyen, Max Petzold, Göran Bondjers, and Henry Ascher**, “Physical growth during the first year of life. A longitudinal study in rural and urban areas of Hanoi, Vietnam,” *BMC pediatrics*, 2012, 12 (1), 1–9.
- Open Philanthropy**, “South Asian Air Quality - writeup,” *Open Philanthropy*, 2022. <https://www.openphilanthropy.org/research/south-asian-air-quality/>, accessed in May 2023.

- Ornaghi, Sara, Simona Fumagalli, Chiara K Guinea Montalvo, Greta Beretta, Francesca Invernizzi, Antonella Nespoli, and Patrizia Vergani**, "Indirect impact of SARS-CoV-2 pandemic on pregnancy and childbirth outcomes: A nine-month long experience from a university center in Lombardy," *International Journal of Gynecology & Obstetrics*, 2022, 156 (3), 466–474.
- Pignotti, Maria Serenella and Sofia Moratti**, "Regulation of treatment of infants at the edge of viability in Italy: the role of the medical profession?," *Journal of medical ethics*, 2010, 36 (12), 795–797.
- Pope, Daniel P, Vinod Mishra, Lisa Thompson, Amna Rehana Siddiqui, Eva A Rehfuss, Martin Weber, and Nigel G Bruce**, "Risk of low birth weight and stillbirth associated with indoor air pollution from solid fuel use in developing countries," *Epidemiologic reviews*, 2010, 32 (1), 70–81.
- Proctor, Robert N**, "The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll," *Tobacco control*, 2012, 21 (2), 87–91.
- Rajšić, Slavica F, Mirjana D Tasić, Velibor T Novaković, and Milica N Tomašević**, "First assessment of the PM10 and PM2. 5 particulate level in the ambient air of Belgrade city," *Environmental Science and Pollution Research*, 2004, 11 (3), 158–164.
- Rich, David Q, Kaibo Liu, Jinliang Zhang, Sally W Thurston, Timothy P Stevens, Ying Pan, Cathleen Kane, Barry Weinberger, Pamela Ohman-Strickland, Tracey J Woodruff et al.**, "Differences in birth weight associated with the 2008 Beijing Olympics air pollution reduction: results from a natural experiment," *Environmental health perspectives*, 2015, 123 (9), 880–887.
- Rosenzweig, Mark R and Christopher Udry**, "External validity in a stochastic world: Evidence from low-income countries," *The Review of Economic Studies*, 2020, 87 (1), 343–381.
- Saliba, Najat A et al.**, "Concentration measurements and chemical composition of PM10-2.5 and PM2. 5 at a coastal site in Beirut, Lebanon," *Atmospheric Environment*, 2004, 38 (4), 523–531.
- Sassi, Sonia, Mohamed Mehdi Abassi, Pierre Traissac, Houda Ben Gharbia, Agnès Gartner, Francis Delpeuch, and Jalila El Ati**, "Intra-household double burden of malnutrition in a North African nutrition transition context: Magnitude and associated factors of child anaemia with mother excess adiposity," *Public health nutrition*, 2019, 22 (1), 44–54.

- Shaddick, Gavin, Matthew L Thomas, Amelia Green, Michael Brauer, Aaron Donkelaar, Rick Burnett, Howard H Chang, Aaron Cohen, Rita Van Dingenen, Carlos Dora et al.,** “Data integration model for air quality: a hierarchical approach to the global estimation of exposures to ambient air pollution,” *Journal of the Royal Statistical Society Series C: Applied Statistics*, 2018, 67 (1), 231–253.
- Shao, Danqing, Yipeng Du, Shuo Liu, Bert Brunekreef, Kees Meliefste, Qian Zhao, Jie Chen, Xiaoming Song, Meng Wang, Juan Wang et al.,** “Cardiorespiratory responses of air filtration: a randomized crossover intervention trial in seniors living in Beijing: Beijing Indoor Air Purifier Study, BIAPSY,” *Science of the Total Environment*, 2017, 603, 541–549.
- Sram, Radim J,** “Impact of air pollution on the health of the population in parts of the Czech Republic,” *International journal of environmental research and public health*, 2020, 17 (18), 6454.
- Sram, Radim J., Ivan Beneš, Blanka Binková, Jan Dejmek, Donald Horstman, František Kotssovéc, David Otto, Sally D. Perreault, Jirí Rubes, Sherry G. Selevan, Ivan Skálík, Robert K. Stevens, and Joellen Lewtas,** “Teplice Program - The Impact of Air Pollution on Human Health,” *Environmental Health Perspectives*, August 1996, 104 (Supplement 4).
- Stein, Zena and Mervyn Susser,** “The Dutch famine, 1944–1945, and the reproductive process. I. Effects on six indices at birth,” *Pediatric Research*, 1975, 9 (2), 70–76.
- , —, Gerhart Saenger, and Francis Marolla,** “Famine and human development: The Dutch hunger winter of 1944-1945.,” 1975.
- Stojanovska, Suzana, Vesna Dimitrovska, and Rut Feuk,** “Republic of Macedonia: Monitoring the situation of children and women,” *Multiple Indicator Cluster Survey 2005-2006*, 2006. https://www.unicef.org/northmacedonia/media/1706/file/MK_CS_MICS_2006_EN.pdf.
- Su, R, W Zhu, Y Wei, C Wang, H Feng, L Lin, M Hod, E Hadar, and H Yang,** “Relationship of maternal birth weight on maternal and neonatal outcomes: a multicenter study in Beijing,” *Journal of Perinatology*, 2016, 36 (12), 1061–1066.
- Taha, Zainab, Aysha Ibrahim Al Dhaheri, Ludmilla Wikkeling-Scott, Ahmed Ali Hassan, and Dimitrios Papandreou,** “Determinants of Delayed Initiation of Breastfeeding: A Cross-Sectional Multi-

center Study in Abu Dhabi, the United Arab Emirates,” *International journal of environmental research and public health*, 2022, 19 (15), 9048.

Tamim, Hala, Souheil El-Chemaly, Anwar Nassar, Ghina Mumtaz, Afamia Kaddour, Tamar Kabakian-Khasholian, Hassan Fakhoury, Khalid Yunis, and National Collaborative Perinatal Neonatal Network (NCPNN), “Incidence and correlates of cesarean section in a capital city of a middle-income country,” 2007.

Thompson, Lisa M, Nigel Bruce, Brenda Eskenazi, Anaite Diaz, Daniel Pope, and Kirk R Smith, “Impact of reduced maternal exposures to wood smoke from an introduced chimney stove on newborn birth weight in rural Guatemala,” *Environmental health perspectives*, 2011, 119 (10), 1489–1494.

Tilstra, Andrea M and Ryan K Masters, “Worth the weight? Recent trends in obstetric practices, gestational age, and birth weight in the United States,” *Demography*, 2020, 57 (1), 99–121.

Tran, Long K, Lidia Morawska, Tran N Quang, Rohan E Jayaratne, Nguyen T Hue, Mac V Dat, Thai Ha Phi, and Phong K Thai, “The impact of incense burning on indoor PM_{2.5} concentrations in residential houses in Hanoi, Vietnam,” *Building and Environment*, 2021, 205, 108228.

Tran, Toan K, BO Eriksson, Chuc TK Nguyen, Peter Horby, Goran Bondjers, and Max Petzold, “Do-daLab: an urban health and demographic surveillance site, the first three years in Hanoi, Vietnam,” *Scandinavian journal of public health*, 2012, 40 (8), 765–772.

United Nations, “Human Development Index (HDI),” *Human development reports*, 2020. Available at https://web.archive.org/web/20220613214005/https://hdr.undp.org/sites/default/files/data/2020/2020_Statistical_Annex_Table_1.xlsx.

US EPA, “Policy Assessment for the Review of the Primary National Ambient Air Quality Standards for Oxides of Nitrogen,” *Office of Air Quality Planning and Standards - Health and Environmental Impacts Division*, 2017, EPA-452/R-17-003. https://www.epa.gov/sites/default/files/2017-04/documents/policy_assessment_for_the_review_of_the_no2_naaqs_-_final_report.pdf.

—, “Technical Support Document (TSD) for the Final Revised Cross-State Air Pollution Rule Update for the 2008 Ozone Season NAAQS: Estimating PM_{2.5}- and Ozone-Attributable Health

Benefits,” March 2021, (Docket ID No. EPA-HQ-OAR-2020-0272). https://web.archive.org/web/20230316230647/https://www.epa.gov/sites/default/files/2021-03/documents/estimating_pm2.5-_and_ozone-attributable_health_benefits_tsd.pdf.

VanderWeele, Tyler J, John D Lantos, and Diane S Lauderdale, “Rising preterm birth rates, 1989–2004: changing demographics or changing obstetric practice?,” *Social science & medicine*, 2012, 74 (2), 196–201.

Vivalt, Eva, “How much can we generalize from impact evaluations?,” *Journal of the European Economic Association*, 2020, 18 (6), 3045–3089.

Vliet, Guy Van, Shiliang Liu, and Michael S Kramer, “Decreasing sex difference in birth weight,” *Epidemiology*, 2009, 20 (4), 622.

Voorheis, John et al., “Air quality, human capital formation and the long-term effects of environmental inequality at birth,” *US Census Center for Economic Studies, CARRA Working Paper Series*, 2017, 5.

Wang, Shanshan, Liren Yang, Li Shang, Wenfang Yang, Cuifang Qi, Liyan Huang, Guilan Xie, Ruiqi Wang, and Mei Chun Chung, “Changing trends of birth weight with maternal age: a cross-sectional study in Xi’an city of Northwestern China,” *BMC Pregnancy and Childbirth*, 2020, 20 (1), 1–8.

WHO, “Ambient air pollution: A global assessment of exposure and burden of disease,” 2016. <https://web.archive.org/web/20220218043604/https://apps.who.int/iris/bitstream/handle/10665/250141/9789241511/9789241511/9789241511353-eng.pdf?sequence=1>.

WHO, “WHO Global Urban Ambient Air Pollution Database (Update 2016),” WHO, 2016. WHO_AAP_database_May2016_v3web.

WHO, “SDG Indicator 11.6.2 Concentrations of fine particulate matter (PM2.5) (at the Urban level),” *World Health Organization Global Health Observatory*, 2023. <https://www.who.int/>

[data/gho/data/indicators/indicator-details/GHO/concentrations-of-fine-particulate-matter-\(pm2-5\)](https://data/gho/data/indicators/indicator-details/GHO/concentrations-of-fine-particulate-matter-(pm2-5)) Accessed: 6/26/2023.

World Bank, “GDP per capita (current US\$),” *World Bank national accounts data, and OECD National Accounts data files*, 2022. <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>, accessed 2/28/2022.

—, “Life expectancy at birth, total (years),” 2022. <https://data.worldbank.org/indicator/SP.DYN.LE00.IN> Accessed: 3/2/2022.

Wu, Xiao, Rachel C Nethery, M Benjamin Sabath, Danielle Braun, and Francesca Dominici, “Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis,” *Science advances*, 2020, 6 (45), eabd4049.

Zhang, Ao, Yumeng Liu, Bin Zhao, Yinping Zhang, Haidong Kan, Zhuohui Zhao, Furong Deng, Chen Huang, Xiangang Zeng, Yuexia Sun et al., “Indoor PM_{2.5} concentrations in China: a concise review of the literature published in the past 40 years,” *Building and Environment*, 2021, 198, 107898.

Zhang, Liqiang, Weiwei Liu, Kun Hou, Jintai Lin, Chenghu Zhou, Xiaohua Tong, Ziyi Wang, Yuebin Wang, Yanxiao Jiang, Ziwei Wang et al., “Air pollution-induced missed abortion risk for pregnancies,” *Nature sustainability*, 2019, 2 (11), 1011–1017.

Zhou, Zehong, Lixue Chen, Hongping Wu, Danni Zheng, Rong Li, Ben Willem Mol, and Jie Qiao, “Assisted reproductive technology in Beijing, 2013–2015,” *Reproductive BioMedicine Online*, 2018, 37 (5), 521–532.

C Data sources

All pollution data comes from [WHO \(2016b\)](#) unless noted below. Preterm means born at gestation < 37 weeks unless noted.

Abu Dhabi, United Arab Emirates

[Taha et al. \(2022\)](#). Mean birth weight: 3,080g (Table 1). N=1,610. Study inclusion: “Mothers with complete data on sociodemographic factors (such as age and education), health factors related to pregnancy and the mode of delivery, and breastfeeding practices (such as breastfeeding initiation) were considered as participants in this study” (p. 3). Their treatment of multiple births is not mentioned. Pollution year: 2013. An alternate pollution estimate closer in time is very similar: [Abuelgasim and Farahat \(2020\)](#).

Astana, Kazakhstan

[Aimukhametova et al. \(2012\)](#). Low birth weight incidence: 5.1% (Table 2). Preterm: 8.3% (Table 2). N = 157. Women with completed singleton pregnancies at a tertiary hospital in Astana who gave birth at 24 weeks or more of gestation (p. 84).

Pollution source [Kenessariyev et al. \(2013\)](#). Annual average PM2.5 estimate: 52.9 (Table 4). Used because Kazakhstan is not included in the [WHO \(2016b\)](#) database.

Beijing, China

[Su et al. \(2016\)](#). Mean birth weight: 3,343g. Low birth weight incidence: 3.6%. (In both cases these are weighted averages taken using outcomes and group sizes in Table 3.) Preterm: 5.0% (Table 2). N=5,479. Sampling: ‘Medical and obstetrical data for each participant was collected from 15 hospitals in Beijing by a systemic cluster sampling survey conducted from 20 June 2013 to 30 November 2013’ (p. 1062). Eligibility was defined as all women “who delivered a live born singleton infant between 20 June 2013 and 30 November 2013 and that were born at 1970 or later” (p. 1062), so mothers could not be older than 43. Pollution year: 2013.

Beirut, Lebanon

[Tamim et al. \(2007\)](#). Low birth weight incidence: 9.18% This comes from counting the births in the low birth weight categories broken out in Table 2. N=18,727. Sampling: All births from nine major hospitals in Beirut and its suburbs in 2001-2002. Pollution year: 2014. An alternate pollution estimate closer in time to the births is very similar: [Saliba et al. \(2004\)](#).

They report a preterm rate of 15.9%, but defined as gestation ≤ 37 weeks instead of < 37 weeks. The US rate would be higher than 15.9% if defined this way.

Belgrade, Serbia

[Maric et al. \(2010\)](#). Mean birth weight: 3,259; N=2,581 (Table 2). Preterm: 8.7% (Table 2; this is defined as ≤ 37 weeks, p. 83). Sample: Births at the Institute of Gynecology and Obstetrics in Belgrade, 1996 to 2003. Only the control group, which was not exposed to the bombings in March-June 1999, is used. Multiple births included (Table 2). Stillbirth rate: 1.3%.

Pollution source [Rajšić et al. \(2004\)](#) Sample period: June 2002 and December 2002. PM2.5: 61. Used to better match the time period covered in the birth data. [WHO \(2016b\)](#) reports PM2.5 of 37 in 2013.

Brescia, Italy

[ATS Brescia \(2020\)](#). Low birth weight incidence: 6.4% in 2003 and 7.2% in 2020. See Figura 21, which includes annual data from 2003 to 2020. Preterm: 6.83% (p. 17). Sample size: 8,045. Sampling: all births to at least one parent living in Brescia. The sampling is described on page 2. Stillbirth rate: 0.33%. Pollution year: 2013.

Colombo, Sri Lanka

[Jayatissa and Hossaine \(2010\)](#). Low birth weight incidence: 16.1%. N=261 (reported in Table 4.3, p. 26). Sampling: randomly-chosen households in Colombo. Pollution year: 2011.

Doha, Qatar

[Abdulkader et al. \(2013\)](#). Low birth weight incidence: 12.36% (Table 1). Preterm: 14% (Table 3). N = 890. Sampling: births at Women's Hospital HMC from March to April 2011. "All babies born live during the study period, including multiples, were included in the study" (p. 33). Pollution year: 2012.

Hanoi, Vietnam

[Nguyen et al. \(2012\)](#). Mean birth weight: 3,251 grams; N=537. Estimate is the weighted average of “urban boy” and “urban girl” birth weight means in Table 1. Survey data on children born at the Dodalab ([Tran et al., 2012](#)) in Dongda, Hanoi, March, 2009 to June 2010. The authors excluded 20 children who were either twins or born with a congenital disease.

Pollution source [Luong et al. \(2017\)](#). Sample period: September 2010 to September 2011. PM 2.5 Estimate: 67. Used because Vietnam is missing from the [WHO \(2016b\)](#) list.

Makkah, Saudi Arabia

[General Authority for Statistics \(2016\)](#). Low birth weight incidence: 9.28%. Calculated using the counts reported in Table 24-1 “Number of Saudi live births during the 5 years preceding the survey, by sex and weight of the child at birth and administrative Area” on page 78. N = 443,128. No sample restrictions reported. Pollution year: 2014.

Monza, Italy

[Ornaghi et al. \(2022\)](#). Average birth weight: 3,170.5 grams. Reported in the text on page 468. Preterm: 6.7% (Table 2). N = 1,882. Hospital data from women giving both at a hospital in Monza. (The 2020 sample excludes women with a symptomatic COVID infection; I use the 2019 estimate.) Stillbirth rate: 0.4%. Pollution year: 2013. Alternative PM2.5 estimate: 31 ([Collivignarelli et al., 2021](#), Figure 4).

Muscat, Oman

[Ministry of Health Sultanate of Oman \(2012\)](#). Low birth weight incidence: 9.95% (Table 8-5: Births in MOH Hospitals and Health Centres according to Governorates during 2012). N=13,576. All births in Royal Hospital, Khawlah Hospital, and Qurayyat Hospital in 2012. Pollution year: 2009.

Skopje, North Macedonia

[Stojanovska et al. \(2006\)](#). Low birth weight incidence: 5.4%. N=213. Reported in Table NU.8: Low birth weight infants (page 72). Sampling: randomly selected households (p. 17). Pollution year: 2013.

Temuco, Chile

[Lucila et al. \(2016\)](#). Average birth weight: 3,386 grams; N=339 (Table 1). Low birth weight incidence:

3.3% (Table 1). Preterm: 6.5% (Table 1). Sample: consenting households with “at least one child under the age of five had resided in the home in the past five years and/or a child had been born alive or stillborn in the household in the last 3 years” (p. 4). Birth weight was collected from health records (p. 4). Pollution year: 2014.

Tunis, Tunisia

[Sassi et al. \(2019\)](#). Mean birth weight: 3,281.7 grams (p. 46). N=437. Survey data from a random sample of “20–49-year-old, non-pregnant women” in 2009-2010. in Greater Tunis. Pollution year: 2010.

Tuzla, Bosnia and Herzegovina

[Halilović and Begić \(2015\)](#). Mean birth weight: 3,387 grams. Low birth weight incidence: 6.24% (p. 345) N=4,096 (p. 345). Hospital data from Infants born in Obstetrics and Gynecology Clinic, University Clinical Center Tuzla in 2007. No noted sample restrictions. Stillbirth rate: 0.7%. Pollution year: 2010.

Ulaanbaatar, Mongolia

[Barn et al. \(2018b\)](#), Table 3). Mean birth weight: 3,490 grams, N=223. (This mean is reported in Section 2.7. Table 3 reports the main outcomes but using median rather than mean birth weight). Sampling: “We recruited women who met the following criteria: ≥ 18 years of age, ≤ 18 weeks of a single-gestation pregnancy, non-smoker, living in an apartment, planning to give birth in a medical facility in the city, and not using a residential portable air cleaner at enrollment. We excluded women who lived in ger households because electricity is unreliable in ger neighborhoods and gers may have higher indoor-outdoor air exchange rates, which reduces HEPA cleaner effectiveness” (p. 982).

Since Ulaanbaatar features prominently in this analysis, I note another pollution source ([Ganbat et al., 2020](#)). Averaging across eight sensors placed throughout the city (Table 1), the authors report a yearly mean PM2.5 of between 50 and 86 in 2014-2019 (Table 3).

United States

I use US natality data from 2019 to calculate birth weight outcomes in the US ([NCHS, 2023](#)). This allows me to exclude non-singleton births.

- **Singletons only:** Mean birth weight: 3,286 grams. Low birth weight incidence: 6.67%. N=3,630,218.

- **All births:** Mean birth weight: 3,254 grams. Low birth weight incidence: 8.31%. N=3,753,815.

Pollution I use a PM2.5 of 10 for the US. This is from the outcome mean in Table 2B column (1) of (Alexander and Schwandt, 2022), which uses EPA data from 2007-2015. This is slightly higher than other estimates. WHO (2023) report an average of 8.4 for the US in 2010-2019. The population-weighted average across counties in Wu et al. (2020) for 2000-2016 is 9.1. Using these lower estimates of PM2.5 would heighten the difference between the predicted and actual birth weights in the polluted places. The US 99th percentile of 15.35 comes from taking the population-weighted 99th percentile of county-year PM2.5 estimates from Wu et al. (2020), weighted by population.

Estimates of the county-level distribution of PM2.5 exposure are population-weighted and use the county-level data on PM2.5 exposure for the years 2000-2016 from Wu et al. (2020), which is based on satellite estimates from Van Donkelaar et al. (2019) (as in Alexander and Schwandt (2022)). For calculating the percent change in low birth weight, I use the mean incidence of 6.7% calculated using 2019 US natality data (NCHS, 2023). The Currie and Walker (2011) estimates are summarized in Alexander and Schwandt (2022, Table A.9).

D Additional sources

Here I detail some sources that did not satisfy the inclusion criteria.

Teplice, 1994-1997

Teplice was one of the most polluted cities in the world in the 1990s, but not currently, so it was not included in [WHO \(2016b\)](#). The Czech government began a collaboration with the US EPA in the early 1990s to study the health consequences of pollution. [Sram et al. \(1996\)](#) (the “Teplice Program”) is a detailed report on pollution exposure and its consequences in Teplice. The city had a mean PM_{2.5} of 122 in January-March 1993 and 28.7 in May-August 1993, and [Dejmek et al. \(1999\)](#) report a mean of 47.6 from May 1993 to March 1996.

At the time of writing, [Sram et al. \(1996\)](#) were in the middle of a study of birth outcomes in Teplice from 1994-1997. They report data for the first 15 months of the study, covering all hospitalized births in Teplice. Out of the 1,626 births in Teplice, the incidence of low birth weight was 8.8% and the incidence of premature birth was 6.2%. The same figures for the 1,380 births to European parents were 6.4% and 5.0% (Table 8). In the US in 2021, the preterm birth rate was 10.5% ([CDC, 2023](#)) and about 8% for singletons ([Martin and Osterman, 2021](#)).⁷ These findings were expanded on in [Dejmek et al. \(1999\)](#), but for that study the authors dropped preterm births. They report a similar fraction of preterm births. Mean birth weight in the sample of term births was 3,354 grams.⁸

Their research also confirmed, using personal sensors from the US EPA, that personal exposure was indeed high in Teplice. In a sample of women with outdoor jobs, PM_{2.5} exposure as measured by the personal sensors ranged from 33 to 106 depending on the month (Table 2 in [Sram, 2020](#)).

Beijing term births, 2007-2009

[Rich et al. \(2015\)](#) study the effect of a large pollution reduction in Beijing due to the Olympics. They analyze birth outcomes in four districts where they deployed specific pollution sensors. Their birth data comes from Beijing hospitals in those districts in 2007-2009 and they begin with 140,298 births. Unfortunately, their study is restricted only to term births (37-41 weeks of gestation) so it cannot be

⁷It’s possible [Sram et al. \(1996\)](#) restricted to singletons although it’s not reported. [Sram et al. \(1996\)](#) also note that, at the time, 42% of women from Teplice smoked in early pregnancy (p. 709). This compares to 7.2% in the US in 2016 ([Drake et al., 2018](#)).

⁸[Dejmek et al. \(1999, p. 477\)](#) dropped 92 preterm births out of 2,045 births or 4.5% (after dropping non-European births, twins, duplicate mothers. Mean birth weight is calculated using the two groups in Table 1.

used as population estimates. But in their methods (p. 881), they describe that this entails dropping about 5% of the sample. This suggests that the rate of premature birth is much lower in Beijing, since recent estimates put the US around 10% ([CDC, 2023](#)). Their sensors measured PM2.5 above 60 for all three years of the sample (Figure 1).

Thailand cities, 2006

[Kosirinond and Son \(2008\)](#) is a United Nations report on Thailand birth weights in 2006 in multiple places, including two of the highly-polluted cities in [WHO \(2016b\)](#): Saraburi: low birth weight incidence: 9.12%, N=8,741. Nakhon Sawan: 9.10%, N=10,548 (“Low birth weight by sex, and province” p. 221). The calculation includes all live births (see the data description on page 218). However, the [WHO \(2016b\)](#) pollution estimates are from 2014, more than five years from these birth weight outcomes. I could not find pollution estimates taken closer to 2006.

E Possible explanations

E.1 Studies of personal exposure

I searched for studies of indoor pollution or personal exposure from the list of cities included in the international sample and from the historical US. The evidence here is limited because, unlike ambient pollution, there are not systematic studies meant to generate estimates of personal exposure at the city or county level, even within the US. In general, however, studies of indoor environments or personal exposure in these places find similarly extreme levels of pollution.

China has some of the best evidence. [Liang et al. \(2019\)](#) had 50 Beijing residents wear portable environmental sensors as they went about their usual routines, recording an average PM2.5 exposure of 38, more than double the worst outdoor measures in the US. A similar study found an average PM2.5 exposure of 127 across 60 office workers in Beijing ([Baccarelli et al., 2014](#)). [Rich et al. \(2015\)](#), a study of air pollution and birth weight outcomes around the Olympics in Beijing, find mean PM 2.5 exposure of 61 when restricting to districts where the mothers in the sample reside.⁹ [Zhang et al. \(2021\)](#) reviews studies of China from 1980-2019 and finds an average indoor PM2.5 of 95 in urban residences. This suggests that high levels of pollution in China result in high levels of exposure.

Apart from China, researchers have studied personal exposure in a handful of other cities from the international sample. [Jorquera et al. \(2018\)](#) find that median indoor PM2.5 across 63 households in urban Temuco, Chile was 44.4. [Chamseddine et al. \(2019\)](#) report high PM2.5 levels across several rooms in a hospital in urban Beirut (Table 3). [Tran et al. \(2021\)](#) record an average indoor PM2.5 of 52.1 and outdoor PM2.5 of 54.4 across 32 urban residential homes in Hanoi.¹⁰ [Borgini et al. \(2015\)](#) had 90 Milan students ages 12 to 18 wear a portable sensor to track their pollution exposure (Milan is not in the sample but is in the same Lombardy region as Monza and Brescia). PM2.5 exposure was 35 indoors and 46 outside (Table 1). There is less evidence from the 1970s United States as the study of indoor pollution was in its infancy.¹¹

⁹This study also finds normal birth weights in Beijing. However, they restricted to term births so this was not used for the main analysis.

¹⁰97% of households had indoor PM2.5 above 25, and 88% had outdoor PM2.5 above 25 (Fig. 4).

¹¹One early example is [Moschandreas et al. \(1981\)](#), which studies the relationship between indoor and outdoor TSP at ten residences and two office buildings in the Boston area. They found that indoor TSP was consistently higher than outdoor TSP.

F Selection effects

F.1 Maternal characteristics in international sample

Unsurprisingly, the highly-polluted international cities are in countries that are poorer than the United States, so selection effects would generally predict worse outcomes. I show country-level life expectancy and GDP for these places in [Table B.2](#). The US is richer than all countries in the sample. And in 2019, US life expectancy was 79, exceeding most counties in the sample but not Italy, Qatar, Lebanon, and Chile. All of these countries rank lower on health and development ([United Nations, 2020](#)). With worse attributes along these dimensions, it would be surprising if some unconsidered factor were bolstering birth weight in these cities.

The mothers in the samples could be different from other mothers in the city in terms of their pollution exposure, health, and demographics. [Currie and Schwandt \(2016\)](#) argue that mothers exposed to the 9/11 dust cloud may have been positively selected along key dimensions, leading to a spurious cross-sectional finding that mothers close to the dust cloud saw better birth outcomes. I noted above that the Ulaanbaatar sample is selected: in particular, it is restricted to non-smokers. Two sources are city-wide: Brescia ([ATS Brescia, 2020](#)) and Makkah ([General Authority for Statistics, 2016](#)). This would seem to address issues around selection within the city. Descriptives are not available for most of the samples included. But the Hanoi sample ([Tran et al., 2012](#), Table 3) shows similar incomes compared to the rest of Vietnam. And 13.6% of the Temuco sample reports smoking during pregnancy, compared to 7.2% in the US ([Drake et al., 2018](#)).

It's possible that inherited factors not related to economic development tend to lead to consistent differences in birth outcomes along racial or ethnic lines. In practice this seems not to threaten the conclusions here as reference materials on birth weight tend to suggest larger weights among European populations ([Janssen et al., 2007](#)). Age is another potential confound. Birth weight increases with maternal age until around 30 ([Wang et al., 2020](#)). Except for China, where the average age at childbirth is similar to the US, this tends to bias comparisons in favor of the US.¹² Taken together, it seems unlikely that background traits of the mothers in the sample predict better birth outcomes overall.

¹²The average age at first birth in China was 26.9 years in 2016 ([He et al., 2019](#), Table 4). In the US it was 26.3 years in 2014 ([Mathews and Hamilton, 2016](#)).

F.2 Maternal characteristics in US sample

Many important maternal characteristics in the US have shifted over this time. This could mask improvements in infant outcomes caused by the drop in pollution from 1972 to 2002. In [Figure A.3](#), I show how trends in birth weight and low birth weight change with the inclusion of controls for mother age, race, live birth order, and state of residence. The trends are unchanged, suggesting that maternal characteristics are not a major confound for the comparisons across time periods. (These results are identical when I restrict to the polluted county sample.)

G Survivorship bias or culling effects

Whether fetal insults have culling or scarring effects is a core question in the fetal origins literature. See [Almond and Currie \(2011\)](#), [Almond \(2006\)](#), and [Chay and Greenstone \(2003\)](#) for discussion and specific examples. [Floris et al. \(2021\)](#) and [Brown and Thomas \(2019\)](#) argue that the flu pandemic of 1918 induced *negative* selection; i.e., infants that survived to birth had lower socioeconomic status. [Bruckner and Catalano \(2007\)](#) find evidence for culling: males born in times of low male:female sex ratios (a proxy for fetal stress) have lower mortality, which suggests that infants who survive to birth in inhospitable conditions have better health outcomes. [Almond and Currie \(2011, p. 157\)](#) discuss this in the context of the null findings from the famines in Finland and Leningrad, where mortality might have been high enough to induce a culling effect. [Bozzoli et al. \(2007\)](#), looking at the relationship between neonatal mortality and adult height, find a negative relationship overall (consistent with scarring), but some evidence of a positive relationship in the poorest countries (consistent with culling).

Some research suggests that the share of male infants falls in response to stressors (e.g., [Catalano et al., 2006](#)). But birth weight is slightly higher for males ([Van Vliet et al., 2009](#)), so effects operating through the sex ratio could decrease, not increase, aggregate birth weight. And in a large study of Ulaanbaatar ([Dorj et al., 2014](#)), the authors record 78,076 female and 82,381 male births, a male-female sex ratio of 51.34%. This is about equal to the sex ratio at birth in the United States ([Mathews et al., 2005](#)), although [Dorj et al. \(2014\)](#) include only full-term births.

In the present context, mothers in polluted and impoverished areas may have higher rates of miscarriage and stillbirth or more difficulty conceiving. If the fetuses that suffer a miscarriage or were not conceived would have had low birth weights, newborn infants in the polluted cities could be positively selected through a culling effect. This question—whether health insults have culling or scarring effects—is prominent in the fetal origins literature. Could selective survival lead to a healthier infant population, at least as measured by birth weight?

This would have the interesting implication that interventions reducing pollution exposure could *decrease* average birth weights through its effects on survival. I could find no studies arguing that fetal insults increase aggregate average birth outcomes through culling, so this explanation would contradict a vast literature (see the meta-analysis in, e.g., [Bekkar et al., 2020](#); [Li et al., 2017](#)). In the developing

world, studies of household fuels tend to find that indoor pollutants both increase stillbirths and decrease birth weight, consistent with a scarring rather than culling effect (Pope et al., 2010; Amegah et al., 2014). Further, while data on miscarriage is sparse and complicated to measure (for the US, see Mukherjee et al., 2013), the rates of stillbirth (Hug et al., 2021) and low birth weight (Blencowe et al., 2019) at a national level tend to move together, which also suggests that deprivation tends to lead to scarring. Still, given the unique sample of extremely polluted cities, the mechanism deserves consideration.

G.1 Culling through miscarriage

To affect birth weight, culling would need to happen either through miscarriage, stillbirth, or (a decrease in) conceptions. Observational studies tend to find that pollution increases spontaneous abortion (Zhang et al., 2019; Grippo et al., 2018). A simpler test is whether the miscarriage/stillbirth rate is in general higher in polluted areas.

The Barn et al. (2018b) study provides a useful estimate of this statistic, which is difficult to find for the countries in the sample. The mothers were recruited at a median gestation of 11 weeks and at or before 18 weeks of gestation.¹³ They report that 10.1% had either a miscarriage or stillbirth. I could find only one other detailed study of miscarriage from a sample with high pollution exposure. Dellicour et al. (2016) recruited 1,134 pregnant women in rural Siaya County, Kenya to estimate weekly miscarriage rates. Their sample included 508 mothers captured before the 12th week of gestation. While pollution exposure is not reported in Dellicour et al. (2016), 93% of rural Kenyans primarily use charcoal and firewood, so the exposure to indoor pollutants is likely high (Kenya Ministry of Energy (2019); also see Dida et al. (2022) for data on a neighboring county). Their survival rates are similar: after 11 weeks of gestation, the risk of miscarriage is 8.8% (Dellicour et al., 2016, Table 2).¹⁴

There are few studies of US samples to compare to, but the existing estimates are comparable to Ulaanbaatar and Kenya. Mukherjee et al. (2013) calculates cumulative miscarriage risk for a sample of mothers in North Carolina. The risk conditional on making it through the 9th week of pregnancy is 8%.¹⁵ Ammon Avalos et al. (2012) collate data from four previous studies in the US and find that mis-

¹³The inter-quartile range (25th-75th percentile) was 9-12 weeks for control and 9-13 weeks for treatment Barn et al. (2018b, Table 1)

¹⁴Stillbirth is unfortunately not reported, but another recent study in Siaya, Eilerts-Spinelli et al. (2022), finds a rate of about 1%.

¹⁵Specifically, "The cumulative risk of embryonic loss (gestational weeks 6–9) was 15% for whites and 14% for blacks. The

carriage rates as of the 9th week of gestation range from 6 to 10% (Figure 3), arguing that the highest estimate is most reliable because it recruited people earlier in pregnancy. The risk of miscarriage as of the 11th week of gestation is 7.7%.

These estimates suggest that miscarriage rates are not distinctly higher among populations with severe pollution exposure, although the comparison is not perfect because [Barn et al. \(2018b\)](#) do not calculate the weekly risk of miscarriage. Assuming all mothers in [Barn et al. \(2018b\)](#) were at exactly 11 weeks of gestation, a bounding exercise shows it is unlikely that this could explain the normal birth weights. Assume the miscarriage rate at 11 weeks is 8% in the US ([Ammon Avalos et al., 2012](#)) and—roughly—10% in Ulaanbaatar ([Barn et al., 2018b](#)). Then there’s a missing 2% of infants in Ulaanbaatar who presumably would have had lower than average birth weights. Assuming these babies are at the 1st percentile of the US birth weight distribution, weighing 1,275g, adding them back into the sample would decrease the Ulaanbaatar average by 43g, from 3,403g to 3,357g, which would still put Ulaanbaatar above the US average. Similarly, the incidence of low birth weight would still be lower in Ulaanbaatar if we counted these infants.

G.2 Culling at conception

A culling effect could also happen earlier, before the infant is conceived. Pollution could inhibit conception for “weaker” embryos. There is some indirect evidence for this: Births in Ulaanbaatar exhibit seasonality, with 13% more births in the Fall compared to the Spring ([Dorji, 2015](#), Table 1).¹⁶ If this drives culling, birth weights should be higher in the season with fewer births, but they are the same ([Dorji, 2015](#), Table 1).

Apart from seasonality, another test is whether couples in highly polluted places have to spend longer trying to conceive. In vitro fertilization (IVF) procedures offer a more controlled environment for measuring this. [Choe et al. \(2018\)](#) find that IVF implantation is slightly less successful for women in more polluted areas. On the other hand, a careful examination by [Zhou et al. \(2018\)](#) finds that implantation success rates are about the same in Beijing as in Europe.

cumulative risk for early fetal loss (weeks 10–15) was 6% and 8% and, for late fetal loss (weeks 16–19), was <1% and 2% for whites and blacks, respectively” ([Mukherjee et al., 2013](#)).

¹⁶[Badarch et al. \(2021\)](#) find greater seasonality, but it’s based on 10,000 births from a single hospital. [Dorji \(2015\)](#) samples 160,000 births from all private hospitals in Ulaanbaatar.

G.3 Obstetric practices

Some research suggests that changing obstetric practices have worked to decrease birth weights in the US: now, c-sections and inductions are more likely to happen before the due date, and pregnancies are not allowed to continue much beyond the 40th week (Tilstra and Masters, 2020; VanderWeele et al., 2012).

These changes could mask improvements in infant health brought by cleaner air in the analysis of US counties from 1972 to 2002. One way to address this is to calculate the birth weight per week and then multiply by 40 to get a constant-gestation birth weight:

$$\text{ConstantGestationBW}_i = 40 * (\text{BirthWeight}_i / \text{Gestation}_i)$$

I show the results in Figure A.2. In contrast to birth weights, constant gestation birth weights have increased—by about 2.6% from 1972 to 2002 in the polluted US county sample. This is small compared to the changes in pollution. The average decrease in pollution was around 60%, so this corresponds with an elasticity of 0.07. The implied elasticity from Alexander and Schwandt (2022) is 1.0 (Figure 6A). The other main outcome used in Figure 1 is low birth weight. This has increased over time in the US sample, and this remains so when I use the constant-gestation correction.

Premature infants are smaller, and the US and other developed countries attain viability—typically defined as a 50% chance the infant survives—at lower lengths of gestation, about 23–24 weeks according to Glass et al. (2015). Some countries in my international sample impose higher thresholds for viability. For example, in China, official guidelines are to only provide full care for infants of at least 28 weeks of gestation (Han et al., 2022). Italy, however, which provides two of the polluted cities from the international sample, almost always provides care at 25 weeks, with debates around the proper practice at 22–25 weeks (Pignotti and Moratti, 2010).

Based on its 2019 data, if the US used China’s same threshold of 28 weeks, average birth weight would be 13 grams more on average, and the incidence of low birth weight would be 6.2% instead of 6.7% (all excluding multiple births). The difference would be even smaller using gestational cutoff of 25 weeks, a conservative guess at Italy’s policy. Data on the thresholds employed by the lower-income countries in the sample is difficult to find. Hayden et al. (2020) surveyed hospitals in the Philippines,

which has a lower GDP per capita than all the countries from the international sample except Vietnam, Iran, and Nigeria. They found that most hospitals provide care at 27-28 weeks of gestation, similar to the official guidelines reported for China. The vast majority of hospitals reported resuscitating at 31-32 weeks.

Even imposing a minimum gestation of 32 weeks on the US sample would leave a large gap between the predicted birth weights and actual birth weights. Restricting to gestation lengths of 32 weeks or more, US birth weights would be about 30 grams higher and 1pp less likely to be low birth weight. (The changes are similar with or without multiple births.) This extreme adjustment would account for just a small share of the gap between the observed birth outcomes and the log predictions discussed above.

G.4 Censoring the US birth distribution

[Figure A.4](#) studies how excluding the bottom 0 to 5 percent of births from the US birth weight distribution would change birth outcomes in the US and the causal predictions in the international sample. As before, multiple births are excluded. Panel (A) shows the results using birth weight. The estimates of the US average birth weight mechanically increase as the amount of censoring increases (solid black line). The red and blue dashed lines show that the average predictions also increase, because they were derived using the US as a reference (see [Section 3.1](#)). Even censoring the bottom 5 percent of the US birth distribution, the polluted city average birth weight, shown in the flat maroon line, exceeds the predictions.

Panel (B) shows the same exercise for low birth weight incidence. The US incidence of low birth weight mechanically decreases with the level of censoring (solid black line). In this case, the average log prediction comes close the observed polluted city average when removing the bottom 5 percent of the US birth weight distribution.

This level of censoring seems implausible given the available evidence. I estimated above that the rate of either stillbirth or miscarriage might be 2 percentage points higher in Ulaanbaatar. This is half the needed censoring, and would not perfectly correlate with birth weight.

H Timing of exposure

The cities and counties in the highly polluted places studied here have stably high levels of pollution. In contrast, much of the quasi-experimental literature is based on shorter-term changes in exposure (as is the case in [Alexander and Schwandt \(2022\)](#) and [Currie and Walker \(2011\)](#), see also the studies reviewed in [Currie et al. \(2014\)](#)). Could it be that short-term effects are larger than long-term effects? This would also be surprising if the primary mechanism is the accumulation of particulate matter in lungs ([Lippmann et al., 2003](#)), although this has not been clearly established.

There are certainly examples from biology where early exposures make individuals more resilient. Human hands and fingers develop calluses in response to repeated friction, which makes them more robust to future exposures. And some research suggests that early exposure to allergens decreases allergic responses ([Hesselmar et al., 1999](#)). In the current context, this explanation could hold that, due to their early exposure, mothers in highly-polluted places are more resilient to pollution than contemporary US mothers. The harmful effects of pollution are thus less likely to impact fetuses in more polluted places.

There is little research available which can speak to this claim. One potential piece of evidence against this is that [Currie et al. \(2009\)](#) find that pollution effects on fetuses were more harmful for mothers who smoked. On the other hand, [Checkley et al. \(2021\)](#) find that wildfire smoke increases premature births much more in places with *less* smoke exposure on average (Figure 4). Overall, it would seem surprising if birth outcomes were worse among infants exposed to a temporary increase in pollution from 5 to 15 PM_{2.5} compared to infants born in an environment with a constant PM_{2.5} of 25. However, it is hard to definitively rule out such a mechanism. Either way, this is more of an explanation for the surprising resilience rather than an argument against it.

H.1 Experimental evidence

A sparse experimental literature tests the idea that short-term changes in air quality could impact fetal health. Based on the quasi-experimental estimates, air purifiers could be a highly cost-effective intervention. Low-cost air purifiers decrease indoor particulate matter by 30-60% ([Barn et al., 2018a](#)) and cost between \$100 and \$200 as of February 2022. People spend most of their time indoors, and outdoor pollution affects indoor pollution ([Leung, 2015](#)). Replacing biomass fuels with ethanol could

have similar benefits.

To my knowledge, there is only one randomized controlled trial of air purifiers and infant health, performed in Ulaanbaatar, one of the most polluted cities in the world ([Barn et al., 2018b](#)). The researchers randomized 540 mothers at 11 weeks gestation. The purifiers caused a 7.2 ppm decrease in indoor PM_{2.5} ([Barn et al., 2018a](#)). However, the effects on birth weight were null with a 95% confidence interval of -84g to 120g. Although this confidence interval encompasses practically large effects, this rejects the linear predicted effects from [Alexander and Schwandt \(2022, Table A.8\(c\)\)](#), which imply a birth weight increase of 167 grams (23.2×7.2). But the wide confidence interval does not rule out my log formulation (see Section 3.1) of the results.

Interestingly, other interventions using air purifiers and liquefied gas achieve large reductions in indoor PM_{2.5} without substantial effects on health outcomes ([Shao et al. \(2017\)](#); [Checkley et al. \(2021\)](#); [Dong et al. \(2019\)](#), also see the review in [GiveWell \(2022\)](#)). These studies use more accurate measures of exposure. In particular, the subjects in [Checkley et al. \(2021\)](#), a trial of liquefied petroleum gas, wear a sensor on their clothing that measures particulate matter in their immediate environment. The sensors record a PM 2.5 of 30 in the treatment group vs. 98 in the control group, but the paper finds no effects on lung function.

Another notable example is [Jack et al. \(2021\)](#), a pre-registered study in rural Ghana with 1,414 households that also provided liquefied petroleum gas to treated families. The gas decreased indoor PM 2.5 by 25 (Figure 3), measured using sensors that were physically placed on the mothers for 72 hours. The birth weights across arms were practically identical. Control infants were 29 grams heavier than treated infants, with the 95% confidence interval rejecting birth weight benefits above 56 grams. The log formulation (see Section 3.1) of the results in [Alexander and Schwandt \(2022\)](#) suggests that this should have increased birth weight by 90 grams. Two similar cooking fuel trials in rural Nepal also found null results on health outcomes ([Katz et al., 2020](#)). As [Jack et al. \(2021\)](#) note, studies of clean fuels have had surprisingly small effects on birth weight.¹⁷ A meta-analysis of the effects in [Jack et al. \(2021\)](#) and [Barn et al. \(2018b\)](#) using a random-effects framework yields a 95% confidence interval of (-0.13, 0.09) standard deviations. Based on birth weight standard deviation of 450 grams, this rejects benefits above 40 grams, while the pooled decrease in PM 2.5 exposure across the two trials was 18.

¹⁷[Thompson et al. \(2011\)](#) estimate a positive impact on birth weight of chimneys but the confidence interval is wide and includes zero.

As [Alexander and Schwandt \(2022\)](#) note, effects from correlational studies of air pollution and health outcomes may be biased downward due to measurement error, but quasi-experimental approaches can correct for this bias. These randomized studies also address the measurement error issue. If pollution sensors are doing a poor job at measuring the biologically-relevant pollution exposure, the random assignment inherent in these trials can be viewed as an instrumental variable similar to the Volkswagen sales. These experimental studies, while unable to reject some large impacts on fetal health, present another puzzle compared to the quasi-experimental evidence.