Prenatal Exposure to PM2.5 and Infant Health : Evidence from Quebec*

Abdel-Hamid Bello† Maripier Isabelle‡ Guy Lacroix§

July 10, 2026

Abstract

We examine the effects of prenatal exposure to fine particulate matter ($PM_{2.5}$) on birth outcomes in a low-pollution setting. We combine administrative birth records for all singleton births in Quebec between 2008 and 2015 with air quality monitoring data and linked tax records. To address residential sorting and seasonal confounding, we exploit within-neighborhood variation using models with neighborhood-by-season fixed effects. We find little evidence that prenatal $PM_{2.5}$ exposure substantially affects birth outcomes at the population level. Estimated effects on birth weight and gestational age are small, and we find essentially no association with low birth weight, preterm birth, or small-for-gestational-age status. However, average effects mask important socioeconomic heterogeneity. We construct a cumulative vulnerability score based on family income, maternal education, and neighborhood deprivation. Adverse effects emerge only among families who cumulate at least two forms of disadvantage. For these children, an increase of 1 $\mu g/m^3$ in average $PM_{2.5}$ exposure is associated with a reduction of approximately 6.6 grams in birth weight. We also find that days

*We thank Caitlin Brown, Bruce Shearer, and Jonathan Zhang for helpful comments and suggestions. We also thank participants at the Société canadienne des sciences économiques annual conference (2025), participants at the CREEi-GRCH-ECOSVieQ student conference at CIRANO (2026), and seminar participants at Université Laval. This research was conducted at the Quebec Interuniversity Centre for Social Statistics (QICSS), part of the Canadian Research Data Centre Network (CRDCN). This service is provided through the support of QICSS' Member Universities, the province of Quebec, the Canada Foundation for Innovation, the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council, the Fonds de Recherche du Québec, and Statistics Canada. The ideas expressed in this text are those of the authors and not necessarily those of the CRDCN, the QICSS or their partners.
Disclosure Statement. We acknowledge financial support from the Fonds de recherche du Québec – Société et Culture (FRQSC) (https://doi.org/10.69777/331610), from the Research Chair on the Evaluation of Public Policies and from the Sentinel North Partnership Research Chair in Economics and Brain Health.

†Université de Montréal, CHU Ste-Justine, GRIP
‡Université Laval, CERVO Brain Research Centre, GRIP, CIRANO
§Université Laval, CIRANO

exceeding the Canadian 24-hour $PM_{2.5}$ standard have disproportionately large effects among vulnerable families.

1 Introduction

Early-life health has lasting consequences. A well-established literature documents that poor birth outcomes predict worse health, educational attainment, and labor market earnings later in life (e.g., Almond et al. 2005; Black et al. 2007; Bharadwaj et al. 2017, 2018), and that reductions in environmental pollution during the prenatal period can generate long-run economic returns (Isen et al., 2017). Among environmental determinants of fetal health, fine particulate matter ($PM_{2.5}$) has received particular attention because of its capacity to affect fetal development through several biological pathways and because it is harder for individuals to avoid than many other pollutants (Cyrys et al., 2004). A large literature documents that prenatal exposure to $PM_{2.5}$ is harmful in high-pollution settings (e.g., Li and Zhang, 2024; Basu et al., 2004). Whether it remains harmful at the much lower concentrations observed in developed countries with strong environmental regulations is less clear. Recent work in low-pollution settings finds limited average effects. For instance, Jahanshahi et al. (2022) find no significant effect on birth outcomes in Northern Ireland, and Ling (2023) reaches a similar conclusion for Norway. These null results raise a natural question: if *in utero* $PM_{2.5}$ exposure does not affect the average newborn in a low-pollution setting, are certain groups of infants still at risk — and if so, do current air quality standards provide adequate protection to all segments of the population?

This paper provides direct evidence on these questions. We combine administrative birth records for approximately 280,000 singleton births in Quebec, Canada between 2008 and 2015 with air quality monitoring data and with mothers' tax and post-secondary education records, linked via social insurance number. The central empirical challenge is residential sorting (Graff Zivin and Neidell, 2013). Lower-income and less educated families tend to live in areas systematically more exposed to higher $PM_{2.5}$ concentrations, so that naive comparisons would conflate the effect of pollution with the effect of socioeconomic disadvantage. We address this by exploiting residual year-to-year variation in $PM_{2.5}$ within neighborhood-by-quarter cells, using models with neighborhood-by-quarter, month, and year fixed effects (Currie et al., 2009). We show that, once this residual variation is isolated, it is far less likely to be correlated with our rich set of socioe-

conomic characteristics. We consider three exposure measures: average $PM_{2.5}$ concentration over the prenatal window, the number of days above the Canadian annual guideline ($10\,\mu g/m^3$), and the number of days above the Canadian 24-hour guideline ($27\,\mu g/m^3$) to understand how, beyond averages, the nature of the exposure affects birth outcomes.

Our results are fourfold. First, consistent with Jahanshahi et al. (2022) and Ling (2023), we find little evidence of average effects of prenatal $PM_{2.5}$ exposure on birth outcomes at the population level. Estimated effects on birth weight and gestational age are small, and we find essentially no association with low birth weight, preterm birth, or small-for-gestational-age status.

Second, average effects mask substantial heterogeneity in newborns' vulnerability to $PM_{2.5}$ exposure based on their socioeconomic background. Notably, we construct a cumulative deprivation score combining three binary indicators: family income below CAD$45,000, residence in a low-income neighborhood, and mother without post-secondary education. We then look separately at families according to their score. Adverse effects are concentrated almost entirely among births with the highest scores, that is, families most likely to accumulate two or three of these disadvantages simultaneously. For these families, we find that a one unit increase in the average prenatal $PM_{2.5}$ exposure can lead to a decrease in birth weight ranging from 3 to more than 6 grams. This gradient is present across all our other exposure measures, and is robust to alternative buffer distances from pollution monitors, spatial clustering corrections, a pre-conception placebo, and a within-mother sibling design. Importantly, this gradient is not primarily driven by differences in ambient exposure: average $PM_{2.5}$ is $9.32\,\mu g/m^3$ among the least deprived families and $10.19\,\mu g/m^3$ among the most deprived, a difference too small to account for the magnitude of the effects we estimate.

Third, we find that this heterogeneity also extends to biological vulnerabilities. Focusing on a sample of newborns admitted to neonatal intensive care units (NICUs), a population with pre-existing vulnerabilities independent of socioeconomic status, we find that a one-unit increase in average $PM_{2.5}$ exposure during the pregnancy is associated with a 32.9-gram reduction in birth weight (approximately twenty-four times the population-level estimate) as well as a 2.3 percentage-point

increase in the probability of low birth weight and a 1.8 percentage-point increase in the probability of preterm birth. In this sample, we also find that our results hold when we account for the important correlation between low birth weight status and prematurity.

Finally, we find that extreme deviations from typical conditions can cause more harm than moderate and sustained increases in average exposure. Among our three exposure measures, days exceeding the Canadian 24-hour guideline of 27 μg/m^3 generate effects that can reach roughly ten times that of an additional day above the annual guideline of 10 μg/m^3, while being about eighteen times less frequent. For vulnerable families, the marginal effect of one additional day above 27 μg/m^3 is comparable in magnitude to a one-unit increase in average $PM_{2.5}$ over the full prenatal window. Strikingly, adverse associations in the NICU sample also emerge for days above the more frequently exceeded annual guideline of 10 μg/m^3. This contrast suggests that biological fragility, like socioeconomic disadvantage, could lower the effective threshold at which prenatal $PM_{2.5}$ exposure can become consequential for infant health.

We contribute to the literature on *in utero* pollution exposure and early health (e.g., Chay and Greenstone, 2003; Currie et al., 2009; Currie and Walker, 2011; Sanders and Stoecker, 2015; Simeonova et al., 2021; Alexander and Schwandt, 2022; Palma et al., 2022; Jahanshahi et al., 2022; Li and Zhang, 2024; Klauber et al., 2024; Gehrsitz, 2017; Pestel and Wozny, 2021) in three ways. First, while existing work in low-pollution settings (e.g., Jahanshahi et al., 2022; Ling, 2023) focuses on average effects, we show that these null averages can conceal meaningful associations for vulnerable subgroups along two dimensions — socioeconomic and biological — and that a simple multidimensional vulnerability score can be sufficient to reveal this pattern.

Second, we document that the nature of exposure matters beyond average levels. As mentioned above, days above the Canadian 24-hour standard of 27 μg/m^3, which occur on only 2% of days in our sample, are associated with effects roughly ten times larger per day than days above the annual standard of 10 μg/m^3. This pattern, consistent across socioeconomic groups, suggests that rare but acute pollution episodes may be particularly consequential for vulnerable populations, even in settings where average annual exposure is well below regulatory thresholds. Results from

our NICU sample also highlight this fact. This has direct implications for the design of air quality standards and alert systems, particularly in an era of increasingly frequent wildfire smoke events. It also suggests that linear projections from estimates based on average exposure may not accurately describe the true costs of *in utero* pollution exposure for infant health.

Third, we provide evidence that jointly modeling the impact of *in utero* pollution exposure on both low birth weight and prematurity (two highly dependent outcomes), while accounting for spatial correlation does not materially affect the conclusions from independent estimations. Our sample of births from a neonatal intensive care unit is particularly well-suited for such an exercise, given that both outcomes are materially higher than in the general sample of births. To our knowledge, such a validation had not been provided in the literature so far.

Our findings are particularly striking given that Quebec operates a universal health coverage system and that the families most affected (often those with low income, limited education, and residence in disadvantaged neighborhoods) are precisely those eligible for targeted prenatal support programs shown to improve birth outcomes (Haeck and Lefebvre, 2016). Neither universal coverage nor targeted nutritional support appears to fully insulate vulnerable families from environmental risks. Our results take on additional significance given evidence that even modest reductions in early-life pollution exposure can generate long-run health and labor market returns (Klauber et al., 2024; Isen et al., 2017), suggesting that the latent costs of prenatal pollution exposure may extend well beyond birth outcomes.

The remainder of the paper is organized as follows. Section 2 describes the biological motivation, the data, and the construction of key variables. Section 3 presents the empirical strategy and discusses identification. Section 4 presents the results. Section 5 concludes.

2 Data

To investigate the relationship between prenatal exposure to $PM_{2.5}$ and health at birth, we combine detailed administrative birth records with air quality monitoring data, tax records, post-secondary

education records, and neighborhood characteristics. Our empirical analysis focuses on several indicators of neonatal health, including birth weight, gestational age, low birth weight, preterm birth, and small-for-gestational-age status. Before describing the data sources, it is useful to briefly review the biological mechanisms through which $PM_{2.5}$ exposure during pregnancy may affect fetal development.

2.1 Biological background

$PM_{2.5}$ consists of airborne particles with aerodynamic diameter below 2.5 micrometers that can penetrate deep into the respiratory system and enter the bloodstream. Exposure during pregnancy has been linked to several biological mechanisms that may adversely affect fetal development. First, $PM_{2.5}$ can induce systemic inflammation and oxidative stress in the mother, potentially impairing placental function and reducing the delivery of oxygen and nutrients to the fetus (Brook et al., 2010; Glinianaia et al., 2004). Second, some components of $PM_{2.5}$, including polycyclic aromatic hydrocarbons and heavy metals, may cross the placental barrier and directly affect fetal growth and development (Jedrychowski et al., 2017). Third, exposure has been associated with maternal conditions such as hypertension and preeclampsia that increase the risk of adverse birth outcomes (van den Hooven et al., 2011).

These mechanisms suggest that prenatal $PM_{2.5}$ exposure may influence fetal growth and pregnancy duration, thereby increasing the risk of outcomes such as low birth weight and preterm birth. More generally, they imply that the relationship between pollution exposure and neonatal health may be non-linear, with adverse effects emerging only beyond certain thresholds or among particularly vulnerable populations. This possibility is especially relevant in low-pollution settings such as Quebec, where average concentrations are generally well below those observed in many previous studies.

The biological pathways linking $PM_{2.5}$ exposure to neonatal health also provide a rationale for the heterogeneous effects investigated in this paper. If pollution affects fetal development through placental function, inflammation, and maternal health conditions, children born into disadvantaged

socioeconomic circumstances may be particularly vulnerable because they face additional risk factors that operate through similar pathways. Likewise, newborns admitted to NICUs may exhibit greater sensitivity to environmental insults because of pre-existing medical fragility. Our analysis therefore examines both average effects and heterogeneity across socioeconomic and biological dimensions.

2.2 Birth record data

Our analysis is based on the Quebec birth certificates included in the Canadian Vital Statistics Database (CVSD) for all singleton births between 2008 and 2015. The database contains detailed information on parental demographic characteristics, including maternal residence during pregnancy, age, place of birth, marital status, and parity, together with neonatal outcomes such as date of birth, sex, birth weight, and gestational age. Our primary outcomes are birth weight (grams), gestational age (weeks), low birth weight (birth weight below 2,500 g), and preterm birth (delivery before 37 completed weeks of gestation). We also examine small-for-gestational-age (SGA) status, defined as birth weight below the 10th percentile for sex and gestational age, computed over all singleton births in our sample.

The birth records are linked to provincial tax files and post-secondary education records, allowing us to control for detailed individual and family socioeconomic characteristics. Prenatal $PM_{2.5}$ exposure is assigned using measurements from the air quality monitor closest to the mother's residence during pregnancy. We further augment the analysis with meteorological variables obtained from nearby monitoring stations and neighborhood characteristics derived from the 2006 and 2016 Canadian Censuses of Population.

To limit the influence of congenital anomalies and birth defects, we restrict the sample to singleton births with birth weight of at least 500 grams and gestational age of at least 26 weeks.[1] This restriction leaves approximately 702,000 singleton births, or about 87,750 births per year.[2]

[1] Information on congenital abnormalities is not recorded in Canadian birth certificates.

[2] This is consistent with the 88,436 births recorded in Quebec in 2010. See https://statistique.quebec.ca/en/document/births-quebec.

We use a confidential version of the CVSD that links each mother to her tax and postsecondary education records for the year of maternity via her social insurance number.[3] This linkage allows us to include individual and family-level socioeconomic variables — including employment status, family income, and education — that have been shown to predict neonatal health outcomes (Kramer, 1987). We successfully link over 97% of observations (about 680,000 births).

Our main regression sample consists of singleton births from mothers living during pregnancy in a Forward Sortation Area (FSA) whose centroid falls within 10 km of an air quality monitoring station operated by the Quebec Ministry of Environment.[4] We calculate the distance between all monitors and each FSA centroid using exact longitude and latitude coordinates, which are available in both our birth and air pollution data. This restriction reduces the sample by approximately 60%, leaving a main regression sample of approximately 282,000 births. The monitors are concentrated in urban areas, so the regression sample overrepresents births in densely populated municipalities.

Table 1 presents summary statistics for both the full and regression samples. Panel A shows that birth outcomes are slightly better in the regression sample: the incidences of low birth weight and prematurity are approximately one percentage point lower on average for children born to mothers living near a pollution monitor. Panel C reveals that mothers in the regression sample are older, more likely to be married, and more likely to be born outside of Canada, reflecting the overrepresentation of immigrant populations in urban areas. University education and income are broadly similar between the two samples. The results are therefore representative of more densely populated areas, where changes in pollution levels could mechanically affect the greatest number of births.

[3]The confidential version of the CVSD is part of the data integration project "The Impact of Preterm Birth on Socioeconomic and Educational Outcomes of Children and Families (IPB)," available at the Canadian Research Data Centre Network.

[4]A FSA is defined by the first three characters of the postal code, . In sensitivity analyses, we also consider alternative buffers of 20 km and 30 km. Results are qualitatively unchanged and are reported in Table B.4.

Table 1: Summary statistics

	Full sample	Residing within 10 km of a monitor
Panel A. Birth outcomes		
Birth weight (g)	3348.13 (547.03)	3316.62 (491.68)
Gestational age (weeks)	38.85 (1.84)	38.68 (1.51)
Low birth weight	0.06 (0.23)	0.05 (0.21)
Preterm birth	0.07 (0.25)	0.06 (0.24)
Small for gestational age	0.10 (0.30)	0.10 (0.30)
Panel B. Child characteristics		
Male	0.51 (0.50)	0.51 (0.50)
First child	0.44 (0.50)	0.45 (0.50)
Panel C. Maternal and family characteristics		
Mother's age	29.74 (5.09)	30.61 (5.15)
Married	0.39 (0.49)	0.51 (0.50)
Immigrant mother	0.22 (0.42)	0.38 (0.48)
University education	0.21 (0.41)	0.23 (0.42)
Low-income neighborhood	0.20 (0.40)	0.26 (0.44)
Family income < $45,000	0.21 (0.46)	0.36 (0.48)
Mother unemployed (t-1)	0.14 (0.33)	0.18 (0.37)
Log family income (t-1)	10.95 (0.85)	10.87 (0.94)
Log mother's income (t-1)	9.94 (1.43)	9.81 (1.67)
Panel D. Pollution exposure		
Average $PM_{2.5}$ ($\mu g/m^3$)	—	9.72 (1.73)
# Days above 10 $\mu g/m^3$	—	98.60 (31.25)
# Days above 27 $\mu g/m^3$	—	5.51 (4.10)
Observations	680,000	282,000

Notes: (1) Each column reports the means and the standard deviations (in parentheses) (2) The full sample includes all singleton births in Quebec between 2008 and 2015; (3) The regression sample is restricted to births whose residential FSA centroid lies within 10 km of an air quality monitoring station; (4) Pollution measures are available only for the regression sample; (5) In line with Statistics Canada data vetting rules, the sample size has been rounded to insure confidentiality. The same applies to all the tables.

2.3 Air Pollution Data

Pollution data are extracted from the Info-Air program of the Quebec Ministry of Environment, which collects air quality readings from 42 stations measuring, among other pollutants, $PM_{2.5}$ concentrations across the province. Because the raw data consist of hourly monitor readings, we aggregate them to daily averages. Figure 1 shows that monitors are concentrated in populated areas, including the cities of Montreal, Laval, and Quebec.

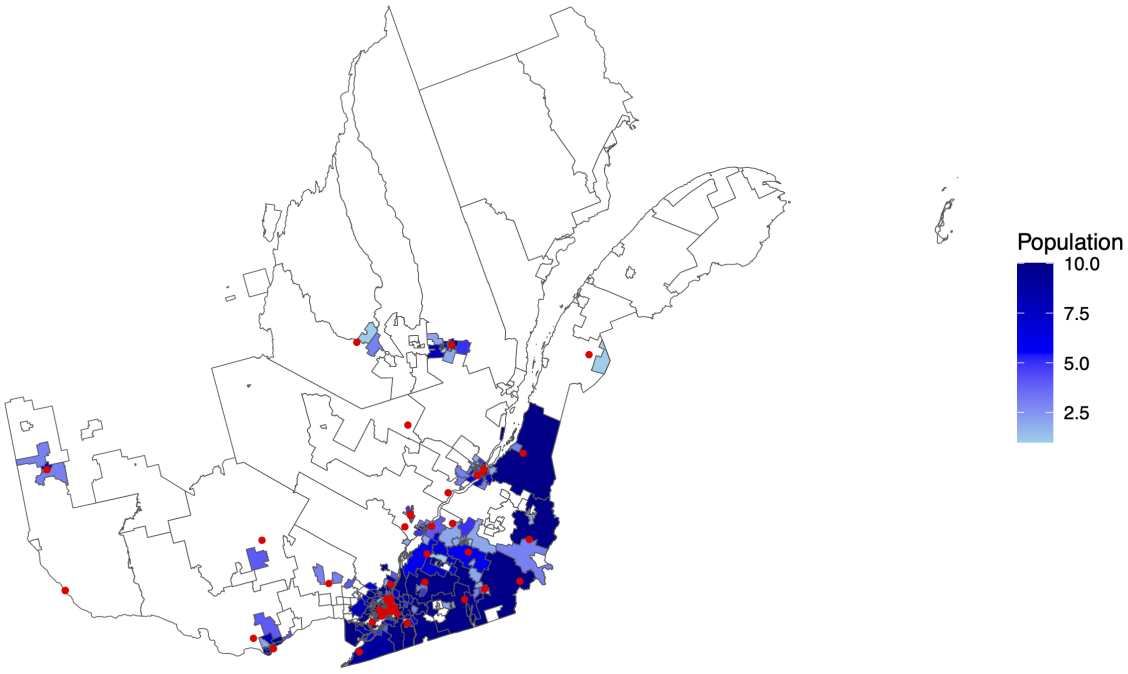


Figure 1: Location of air quality monitoring stations and population density in Quebec

Notes: The figure displays the location of the 42 air quality monitoring stations used in the analysis together with population density from the 2011 Canadian Census. Darker areas indicate more densely populated regions, while white areas correspond to regions for which population density information is unavailable.

Appendix Figure A.1 presents annual and daily PM2.5 concentrations over the study period. Panel (a) shows a modest increase in average concentrations between 2007 and 2011, followed by a gradual decline thereafter. Panel (b) highlights the relatively low levels of ambient pollution observed in Quebec. Average daily PM2.5 concentrations are centered around 7–10μg/m^3, well below the U.S. 24-hour standard of 35μg/m^3 and the Chinese 24-hour standard of 75μg/m^3. The Canadian 24-hour standard of 27μg/m^3 is exceeded only on a small number of days, indicating that episodes of elevated pollution are relatively uncommon during our study period.

Construction of Exposure Measure.—We assign pollution exposure to each mother based on readings from air quality monitors within a 10-km buffer of her FSA centroid.[5] Daily exposure is computed as a weighted average of readings from all monitors within the buffer, weighting each monitor by the inverse of its distance to the FSA's centroid. *In utero* exposure is then calculated by averaging these daily readings over the 38 weeks preceding the child's exact birth date. This is a standard approach in the literature to account for the endogeneity of gestational age: using the actual gestation length would mechanically correlate shorter exposure windows with premature births (e.g., Palma et al., 2022).

This centroid-based approach introduces measurement error relative to assigning exposure based on exact residential coordinates (Graff Zivin and Neidell, 2013). Measurement error is likely to attenuate our estimates toward zero, making them conservative. We also assess the sensitivity to the choice of 20-km and 30-km alternative buffer zones.[6]

In addition to average $PM_{2.5}$ concentration, we construct two threshold-based exposure measures: (1) the number of days during pregnancy on which daily $PM_{2.5}$ exceeded the Canadian annual standard ($10\,\mu g/m^3$), capturing the *frequency* of exposure above this regulatory benchmark; and (2) the number of days with an average concentration above the Canadian 24-hour standard ($27\,\mu g/m^3$), capturing exposure to acute pollution spikes. Together, these three measures allow us to distinguish between the health effects of chronic exposure to lower $PM_{2.5}$ concentrations and short-term pollution peaks.

2.4 Weather Data

Weather data are extracted from Environment Canada's weather information gateway.[7] We limit our search to Quebec stations over 2007–2015 and match each station to the nearest air quality

[5] Our birth data and air pollution data are therefore at the FSA level.

[6] The estimated effects are qualitatively unchanged and slightly smaller under wider buffers, consistent with greater attenuation bias from spatial measurement error at larger distances.

[7] https://climate.weather.gc.ca/historical_data/search_historic_data_e.html.

monitor within 30 km.[8] We retain variables that are systematically reported across all stations: wind direction, wind speed, temperature (°C), and relative humidity. These variables are averaged over the 38-week exposure window to construct pregnancy-level weather controls, with quadratic terms included to capture potential non-linearities in the relationship between weather and health.[9]

2.5 Socioeconomic and Neighborhood Data

Individual and family socioeconomic characteristics are drawn from tax records linked to the CVSD. Our key SES variables are: an indicator for whether the mother was employed in the year prior to birth, an indicator for family income below $45,000 in the year prior to birth, an indicator for whether the mother was married at the time of birth, and an indicator for whether the mother was born outside Canada. University education is derived from postsecondary education records. Neighborhood-level characteristics are drawn from the 2006 and 2016 Canadian Censuses: They include FSA population size, share of immigrants, share of homeowners, and median property value.[10]

3 Econometric Strategy

We estimate the effect of *in utero* $PM_{2.5}$ exposure on infant health outcomes using a two-way fixed effects specification similar to that in Currie et al. (2009):

$$y_{inyqmd} = \beta\, \mathrm{PM}_{2.5_{nd}} + X_i'\,\delta + \gamma \mathrm{Weather}_{nd} + Z_n'\,\theta + \phi_{nq} + \lambda_m + \mu_y + \varepsilon_{inqm}, \tag{1}$$

[8] A finer proximity threshold would result in a significant loss of observations, while the gain in accuracy would be negligible given that weather conditions vary little over short distances.

[9] While a growing literature turns to an instrumental variable approach using, for example, thermal inversions (e.g., You and Chou, 2025; Jans et al., 2018; Chen et al., 2026; Ghafarianghadim and Heutel, 2026) or wind directions (e.g., Deryugina et al., 2019; Graff Zivin et al., 2023), the granularity of data on such variables in Quebec for our sample period does not allow us to employ such an identification strategy.

[10] We use the 2006 Census for births between 2008 and 2010 and the 2016 Census for births between 2011 and 2015. We exclude the 2011 Census due to a major change in sampling methodology.

where i indexes the infant, n indexes the Forward Sortation Area (FSA) of residence, q indexes the quarter of birth, m indexes the month of birth, and y indexes the year of birth. The variable y_{inqmd} denotes one of our health outcomes (birth weight, gestational age, low birth weight, preterm birth, or small for gestational age). $PM_{2.5_{nm}}$ is one of our three exposure measures — average $PM_{2.5}$ concentration, days above 10 μg/m^3, or days above 27 μg/m^3 — averaged or summed over the 38-week exposure window.[11] X_i is a vector of individual and family characteristics: dummy indicators for male infant and parity, maternal age dummies (15–19, 20–24, 35–39, and 40+, with 25–34 as the reference category), and indicators for married mother, and mother born in Canada.[12] Our most complete specifications additionally include an indicator for university education, the log of family income in the year prior to birth, the log of the mother's own income in the year prior to birth, and an indicator for the mother being unemployed in the year prior to birth.[13] Weather$_n m$ is a vector of weather controls (temperature, humidity, wind speed, with quadratic terms, and wind direction dummies) averaged over the exposure window. Z_n is a vector of FSA-level controls from the 2006 and 2016 Censuses. The vector of fixed effects $\phi_{nq} + \lambda_m + \mu_y$ follows exactly the structure in Currie et al. (2009): ϕ_{nq} denotes FSA-by-quarter fixed effects, λ_m denotes month fixed effects, and μ_y denotes year fixed effects. This structure controls for (i) neighborhood-specific seasonal patterns in pollution that might be correlated with seasonal patterns in birth timing, (ii) province-wide monthly shocks common to all neighborhoods in a given month, and (iii) province-wide annual trends, such as the secular decline in $PM_{2.5}$ concentrations after 2011 (Appendix Figure A.1). Standard errors are clustered at the FSA level to account for serial correlation in pollution exposure

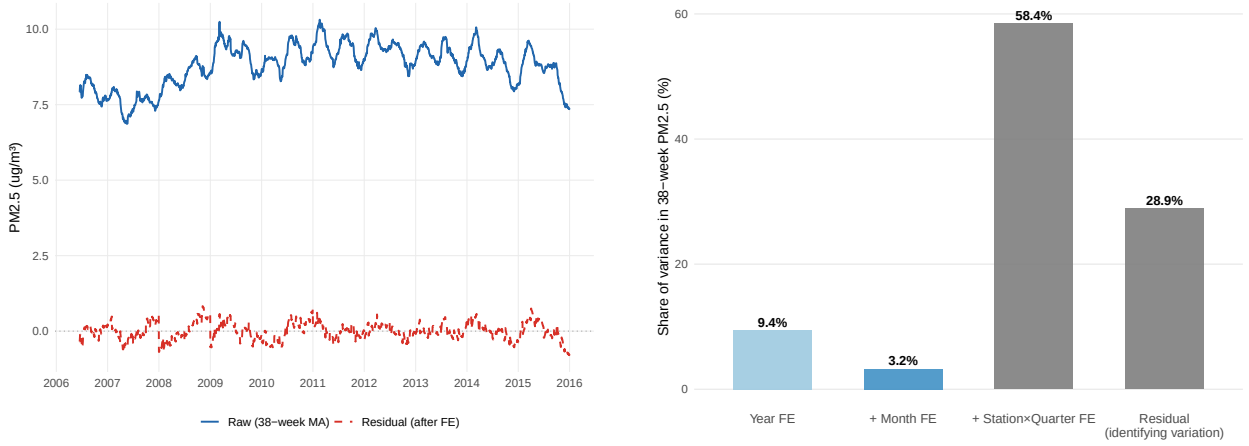
[11] We use precise birth date information to calculate our pollution and weather measures in the 38 weeks preceding the birth.

[12] We also estimated a trimester-specific specification. Because these results are very similar to those obtained from our main yearly specification, and to conserve space, they are not reported but are available upon request. They provide little evidence that the effects of PM2.5 are concentrated in a single trimester. Rather, exposure in each trimester is associated with increases of similar magnitude in the probability of low birth weight. This pattern contrasts with evidence from higher-pollution settings, where specific critical windows of exposure have been identified. In the Quebec context, where average pollution levels are relatively low, the results are more consistent with a cumulative exposure mechanism, whereby PM2.5 affects fetal development throughout pregnancy rather than during a single critical period.

[13] We refer to the specification including these additional socioeconomic variables as the "full" specification, and to the specification excluding them as "full, no SES" in Table 3. Most prior studies proxy income with census-tract median income (e.g., Currie et al., 2009)

within neighborhoods.[14] Figure 2 illustrates the variation used to identify β in equation (1). Panel (a) compares the 38-week moving average of raw $PM_{2.5}$ exposure with its residual after netting out year, month, and FSA-by-quarter fixed effects. While these fixed effects absorb a substantial share of exposure variation, the residual series continues to display meaningful fluctuations over time. Panel (b) formally decomposes the variance of prenatal $PM_{2.5}$ exposure across the different sets of fixed effects. Together, the three fixed-effect dimensions explain approximately 71% of total variation, leaving nearly 29% of exposure variation available for identification.

The identifying assumption is that, conditional on ϕ_{nq}, λ_m, μ_y, and our weather and socioeconomic controls, the remaining variation in $PM_{2.5}$ is uncorrelated with unobserved determinants of infant health. In practice, this identifying variation corresponds to year-to-year idiosyncratic deviations in $PM_{2.5}$ within a given FSA-quarter cell, net of common monthly and annual shock— for example, a year in which a particular neighborhood experiences an unusually smoky summer due to wildfire smoke transport, independent of the neighborhood's average seasonal pollution pattern and of province-wide conditions in that year.



(a) Raw and residual prenatal $PM_{2.5}$ exposure

(b) Variance decomposition

Figure 2: Identifying variation in prenatal $PM_{2.5}$ exposure.

[14] In Section 4.4 we report Conley (1999) standard errors that additionally account for spatial correlation across neighboring FSAs.

Residential sorting and identification. A central threat to identification is residential sorting: families with different socioeconomic characteristics may systematically locate in neighborhoods with different levels of ambient pollution, generating a spurious correlation between $PM_{2.5}$ and infant health that reflects neighborhood selection rather than a causal effect of pollution. Table 2 investigates this issue and illustrates how our fixed effects address it. Panel A ("Before") regresses average *in utero* $PM_{2.5}$ on maternal and household characteristics with no fixed effects or controls: the joint F-statistic of 1,356 confirms strong sorting on socioeconomic status, with higher family income associated with lower exposure (-0.070, $p<0.01$) and residence in a low-income neighborhood or having less than a university education associated with higher exposure ($+0.395$ and $+0.500$, $p<0.01$). Panel B ("After") reports the same regression using the residual of $PM_{2.5}$ once ϕ_{nq}, λ_m, μ_y, weather controls, and FSA-level controls have been absorbed: the joint F-statistic falls to 2.18, a reduction of more than 99.8%, and the few coefficients that remain significant (born outside Canada, mother's income, mother unemployed in $t-1$) are below 0.017 on a sample mean of 9.717 μg/m^3, thus economically negligible. While this does not rule out selection on unobservables, the dramatic reduction in the joint F-statistic substantially mitigates concerns that our estimates are driven by sorting on the characteristics most plausibly correlated with both pollution exposure and infant health.

Table 2: Residential sorting and balance checks

Characteristic	Before FE $PM_{2.5}$	After FE Residual $PM_{2.5}$
Male infant	0.006	0.001
	(0.006)	(0.004)
First child	0.200 ***	0.006
	(0.010)	(0.006)
Second child	0.081 ***	0.000
	(0.009)	(0.006)
Age 15–19	0.022	0.028 *
	(0.027)	(0.016)
Age 20–24	−0.054 ***	0.002
	(0.012)	(0.007)
Age 30–34	0.157 ***	0.001
	(0.008)	(0.005)
Age 35–39	0.266 ***	0.010 *
	(0.010)	(0.006)
Age 40+	0.309 ***	0.001
	(0.016)	(0.001)
Married	0.246 ***	−0.011 ***
	(0.007)	(0.004)
Born outside Canada	0.500 ***	0.017 ***
	(0.008)	(0.005)
University	0.023 ***	0.001
	(0.007)	(0.005)
Log family income	−0.070 ***	0.001
	(0.004)	(0.005)
Log mother income	0.008 ***	−0.003 ***
	(0.002)	(0.001)
Unemployed (year prior)	0.073 ***	−0.014 **
	(0.011)	(0.007)
Poor neighborhood	0.395 ***	−0.005
	(0.007)	(0.004)
Joint F-statistic	1,356.0***	2.18*

Notes: (1) "Before FE" reports coefficients from a regression of average *in utero* $PM_{2.5}$ on maternal and household characteristics, with no fixed effects or controls. "After FE" reports coefficients from the same regression using residual $PM_{2.5}$ after absorbing neighborhood-by-quarter, month, and year fixed effects, weather controls, and neighborhood-level controls. Reference categories: female infant, third or higher birth order, age 25–29, not married, born in Canada, no low-income neighborhood. (2) SE clustered at the neighborhood level. (3) $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

4 Results

4.1 Baseline Results

Table 3 shows little evidence of a strong association between prenatal $PM_{2.5}$ exposure and infant health, irrespective of how exposure is measured. We do not report results for low birth weight, preterm birth, or small for gestational age, as the estimated coefficients are essentially zero across all specifications and exposure measures.[15] The table also shows that models that do not account for FSA×quarter fixed effects may be misleading. In Panel A column 1, the coefficient on average $PM_{2.5}$ suggests that higher exposure is associated with *longer* gestation — the opposite of what one would expect. Once we control for neighborhood-specific seasonal patterns in column (2), the association is in the expected direction, but the coefficient is very small: a one-unit increase in average $PM_{2.5}$ is associated with a 0.004-week (approximately a third of a day) decrease in gestational age, or about 0.01% of the sample mean. The same pattern holds for days above the 27 μg/m^3 threshold (Panel C), where the coefficient is positive and insignificant in column (1) but negative and significant once FSA×quarter fixed effects are included. No statistically significant association is estimated once all fixed effects are included when looking at days above the 10 μg/m^3 threshold (Panel B).

For birth weight, all specifications suggest a negative association between $PM_{2.5}$ and birth weight. Once again, the coefficients are small and become imprecise once we control for income and education (column 3).

[15] Results for these outcomes are available upon request.

Table 3: Population-level average effects of PM2.5 *in utero* exposure on infant health outcomes

	No FSA×Q FE (1)		Full, no SES (2)		Full (3)	
Panel A. Average $PM_{2.5}$ ($\mu g/m^3$)						
Birth weight (g)	-1.941**	(0.898)	-1.690*	(0.878)	-1.398	(0.875)
Gestational age	0.012***	(0.003)	-0.004**	(0.002)	-0.004**	(0.002)
Panel B. Days above 10 $\mu g/m^3$						
Birth weight (g)	-0.106**	(0.049)	-0.092**	(0.046)	-0.078*	(0.046)
Gestational age	0.001***	(0.000)	0.000	(0.000)	0.000	(0.000)
Panel C. Days above 27 $\mu g/m^3$						
Birth weight (g)	-0.927**	(0.383)	-0.742*	(0.410)	-0.485	(0.397)
Gestational age	0.002	(0.002)	-0.001***	(0.000)	-0.001***	(0.000)
N	≈282,000					
Month & year FE	Yes		Yes		Yes	
FSA×quarter FE	No		Yes		Yes	
Weather & FSA controls	Yes		Yes		Yes	
Demographic controls	Yes		Yes		Yes	
SES controls (tax data)	No		No		Yes	

Notes: (1) Mean birth weight = 3,316 g. Mean gestational age = 38.68 weeks. Coef (SE); (2) Standard errors clustered at FSA level; (3) SGA not reported — coefficients are zero across all specifications; (4) Results for LBW and PTB also not shown as parameter estimates are essentially zero (see Section 4.2); (5) *p<0.10; **p<0.05; ***p<0.01. Standard errors are clustered at FSA level level.

4.2 Heterogeneous Effects

Although we find little evidence that *in utero* exposure to $PM_{2.5}$ dramatically affects birth outcomes in Quebec, this null effect may conceal substantial heterogeneity across subgroups. This is particularly relevant in our setting, as differences in biological vulnerability or avoidance behavior across groups may bias the population-level estimate toward zero. We investigate this possibility by augmenting equation (1) with interactions between our exposure measures and a set of indicators for group membership. We first consider heterogeneity by infant sex: a large literature documents that male infants are more vulnerable to *in utero* shocks (e.g., Kraemer, 2000), a pattern known as the

fragile male hypothesis. We then consider heterogeneity along several socioeconomic margins: maternal education (no postsecondary education), family income (below $45,000), neighborhood income (residence in a low-income neighborhood), and immigration status (mother born outside Canada). Table 4 reports the interaction coefficients for birth weight (Panel A) and gestational age (Panel B). Each column corresponds to one of our three exposure measures, and each row reports results from a separate regression interacting that exposure measure with the corresponding subgroup indicator.

We find little evidence of heterogeneity by maternal immigration status and infant gender. None of the interaction coefficients for immigrant mothers are statistically different from zero, across exposure measures or outcomes.

The clearest pattern in Table 4 is that differential effects are concentrated among socioeconomically disadvantaged groups — families with no postsecondary education, low income, and those residing in poor neighborhoods. For these groups, the interaction coefficients are negative and statistically significant across most exposure measures, and the magnitudes are economically meaningful: for instance, an additional day above 27 μg/m^3 is associated with a 2.5-gram additional decrease in birth weight for infants of mothers without postsecondary education, and a 2.2-gram additional decrease for low-income families, relative to their respective reference groups. Some differential effects for low-income families and those in poor neighborhoods are also observed when looking at gestational age, with both average $PM_{2.5}$ and days above 27 μg/m^3 as our exposure measures.

The magnitude of the differential effects is also informative across exposure measures, and the pattern is remarkably consistent across the two socioeconomic margins. Among individuals with both low income and no post-secondary education, the interaction coefficients for average $PM_{2.5}$ and for days above 27 μg/m^3 are of similar magnitude — -2.362 and -2.461 for no post-secondary education, and -2.337 and -2.163 for low income — while the interaction with days above 10 μg/m^3 is roughly an order of magnitude smaller in both cases (-0.215 and -0.194, respectively). Two features of our exposure measures help explain this pattern. First, the two

thresholds differ sharply in frequency: days above 10 μg/m^3 occur on average 98.6 times during the 266-day exposure window — about 37% of days, an almost routine occurrence — whereas days above 27 μg/m^3 occur on average about 5.5 times, or roughly 2% of days. A marginal additional day above the lower, frequently-exceeded threshold therefore represents a comparatively unremarkable change, consistent with its small coefficient. Second, the similarity between the average $PM_{2.5}$ and days-above-27 coefficients is consistent with the two measures capturing exposure shocks of comparable magnitude.[16]

Table 4: Heterogeneous effects of prenatal $PM_{2.5}$ exposure

	Panel A. Birth weight (g)			Panel B. Gestational age (weeks)		
Interaction variable	Avg. $PM_{2.5}$	Days >10	Days >27	Avg. $PM_{2.5}$	Days >10	Days >27
No post-secondary	-2.362***	-0.215***	-2.461***	0.002	0.000	-0.002*
education	(0.369)	(0.034)	(0.403)	(0.003)	(0.000)	(0.001)
Low income	-2.337***	-0.194***	-2.163***	-0.005***	0.000	-0.005***
(<$45,000)	(0.300)	(0.028)	(0.360)	(0.001)	(0.000)	(0.001)
Male infant	-0.629	-0.004	0.024	-0.002*	0.000	0.000
	(1.899)	(0.066)	(0.496)	(0.001)	(0.000)	(0.000)
Immigrant	0.370	0.037	-0.019	-0.004	0.000	-0.002
mother	(1.672)	(0.091)	(0.570)	(0.004)	(0.000)	(0.002)
Poor	-0.948***	-0.080**	-0.805*	-0.003***	0.000	-0.003***
neighborhood	(0.346)	(0.033)	(0.467)	(0.001)	(0.000)	(0.001)
N	≈280,000					

Notes: (1) Interaction coefficients (differential effect for subgroup G=1 relative to reference G=0) (2) All specifications include demographic controls, weather controls, FSA-level controls, month FE, year FE, and FSA×quarter FE following Currie et al. (2009); (3) SE clustered at FSA level in parentheses; (4) *p<0.10; **p<0.05; ***p<0.01.

[16] As a rough illustration, if the relationship between the two exposure measures were proportional, one additional day above 27 μg/m^3 (from 5.51 to 6.51 days) would correspond to an increase in average $PM_{2.5}$ exposure of roughly $9.717 \times (6.51/5.51) \approx 11.48\,\mu$g/m^3, i.e., an increase of about 1.76 μg/m^3 — broadly comparable to the one-unit increase considered in our average $PM_{2.5}$ specification. This calculation is illustrative only, as the relationship between the two exposure measures depends on the full shape of the pollution distribution.

4.3 Cumulative Vulnerability

The results presented above suggest that the adverse effects of prenatal $PM_{2.5}$ exposure, specifically on birth weight, are concentrated among socioeconomically disadvantaged populations. Importantly, the same pattern emerges across several dimensions of disadvantage, including low family income, low educational attainment, and residence in a poor neighborhood. While each of these indicators captures a distinct aspect of socioeconomic status, they are likely to reflect a common underlying vulnerability to environmental risks. However, these characteristics rarely occur in isolation. Families experiencing one form of disadvantage are often exposed to multiple sources of socioeconomic vulnerability simultaneously. As a result, examining each dimension separately may underestimate the extent to which pollution disproportionately affects the most vulnerable children.

To capture the multidimensional nature of disadvantage, we construct a cumulative deprivation score defined as the sum of three binary indicators: (1) family income below $45,000; (2) residence in a low-income neighborhood, defined as the lowest quartile of FSA median income, and (3) no post-secondary education. The resulting score ranges from 0, for families satisfying none of these conditions, to 3, for families satisfying all three. In our sample, the mean score is 1.19, with a standard deviation of 0.860. We then estimate equation (1) for birth weight, interacting each exposure measure with indicators for score values 1, 2, and 3, with score 0 as the reference group.[17] The coefficient for score 0 captures the association for the least deprived families, while the coefficients for scores 1–3 capture differential associations relative to score 0.

Figure 3 and Appendix Table B.1 reveal a striking gradient. For reference families (scoring 0 in our deprivation index), the estimated association between average $PM_{2.5}$ and birth weight is small and statistically insignificant (-0.283 gram, SE = 0.947), and there is virtually no change in that association for families with only one marker of deprivation (0.454 gram, SE = 0.297). Adverse effects emerge only once families cumulate at least two of the three disadvantage indicators. The

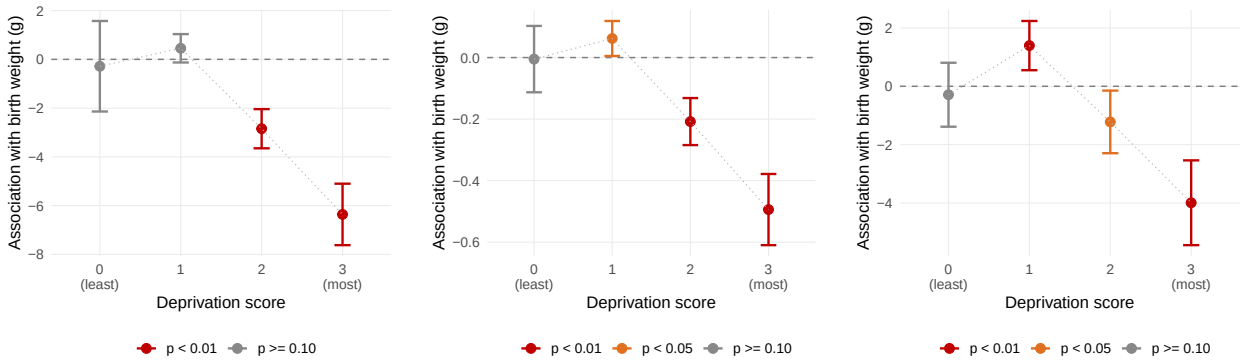
[17] We do not present the results for gestational age as they are very small and/or not statistically significant, as hinted in Table 4.

differential for score = 2 is -2.844 grams (SE = 0.410), yielding a total association of -3.127 grams per unit increase in average $PM_{2.5}$. For score = 3, the differential reaches -6.359 grams (SE = 0.643), for a total association of -6.642 grams — more than 23 times the point estimate for newborns in families without any marker of vulnerability (score = 0). A similar pattern emerges for days above the 10 μg/m^3 threshold. When looking at the association between birth weight and days above the 27 μg/m^3, we again find no significant association for families with score = 0 and the total effect of a day above the threshold is also not statistically significant from 0 for families with a score = 1 on our deprivation index. A differential of -1.220 grams (SE = 0.547) is estimated for a score = 2 (total association of -1.509), and -3.990 grams (SE = 0.741) for a score = 3 (total association of -4.279 grams). The results therefore suggest that vulnerability to prenatal pollution exposure is not confined to the most extremely disadvantaged families: any combination of two or more forms of socioeconomic disadvantage is associated with meaningfully larger effects on birth weight.

Notably, this gradient in associations does not mirror ambient exposure: average $PM_{2.5}$ is 9.32, 10.12, 8.95, and 10.19 μg/m^3 for scores 0 through 3, with no monotonic relationship to the deprivation score, and families with score = 0 and score = 3 face similar ambient levels (9.32 versus 10.19). Yet the total association for score = 3 is more than 23 times larger than for score = 0. This is inconsistent with an explanation based on differential ambient exposure alone, and points to differences across families in how ambient exposure translates into realized biological dose or in underlying susceptibility to a given dose. Several complementary mechanisms could generate this pattern. Differential avoidance behavior — for instance, unequal access to air filtration, time spent outdoors, or occupational exposure — is one candidate, particularly given that avoidance is itself likely to be income-constrained. Differential biological susceptibility linked to co-occurring stressors (nutritional status, chronic stress, or reduced access to prenatal care) offers a complementary explanation, one we return to in Section 4.5 in the context of biologically vulnerable newborns. It is also possible that our exposure measure, based on monitor readings assigned via FSA centroids, is a noisier proxy for true individual exposure among more mobile or more residentially unstable

families, which would attenuate estimates for advantaged and disadvantaged families differently. Our data do not allow us to fully adjudicate between these channels. We note, however, that the gradient is not an artifact of conditioning on socioeconomic controls: Table 3 shows that our main estimates change only modestly when SES controls are added to (or omitted from) the specification, suggesting that these variables are not simply absorbing the very variation that generates the deprivation-score pattern documented here.

These findings suggest that the burden of air pollution in Quebec is not evenly distributed across the population. While average associations are modest, the remaining health burden is concentrated among children facing multiple and overlapping forms of socioeconomic disadvantage. More broadly, the results indicate that cumulative vulnerability plays an important role in shaping susceptibility to environmental exposures, even in a relatively low-pollution setting.



(a) Avg. $PM_{2.5}$ (b) Days $> 10\,\mu g/m^3$ (c) Days $> 27\,\mu g/m^3$

Figure 3: Associations between prenatal $PM_{2.5}$ exposure and birth weight by deprivation score

Notes: Notes: Each point reports a coefficient from equation (1), which interacts exposure with indicators for deprivation scores of 1, 2, and 3 (score 0 is the reference category). The coefficient for score 0 captures the baseline association, while the coefficients for scores 1–3 represent differential associations relative to score 0. Bars denote 95% confidence intervals. Standard errors are clustered at the neighborhood level. The deprivation score is defined as the sum of three indicators: family income below $45,000, residence in a low-income neighborhood (lowest quartile of FSA median income), and no post-secondary education. $p < 0.10$; $p < 0.05$; $p < 0.01$.

4.4 Robustness and Sensitivity Analyses

Because the cumulative vulnerability results reported in Section 4.3 constitute our central finding, we subject the estimates to a series of additional robustness checks addressing three distinct threats

to identification: residual confounding, spatial dependence, and exposure measurement error.

Multiple hypothesis testing. Because our cumulative vulnerability results in Section 4.3 rely on testing differential associations across several deprivation-score levels and exposure measures, we assess whether our findings survive correction for multiple comparisons.[18] We treat the nine hypotheses tested for birth weight — the differential associations for scores 1, 2, and 3 relative to score = 0, estimated separately for average $PM_{2.5}$, days above 10 μg/m^3, and days above 27 μg/m^3 — as a single family, and report both Bonferroni (Holm, 1979) and Benjamini and Hochberg (1995) false discovery rate (FDR) corrected p-values in Appendix Table B.5. The differential associations for score = 2 and score = 3 remain statistically significant at conventional levels under both corrections and across all three exposure measures, indicating that our central finding — that adverse effects are concentrated among births with two or more markers of deprivation — is not an artifact of testing multiple hypotheses. The differential association for score = 1, which is small and only occasionally significant even before correction, does not survive the more conservative Bonferroni adjustment for average $PM_{2.5}$ and days above 10 μg/m^3, though it remains significant under FDR for all three measures. This pattern is consistent with our main interpretation: meaningful adverse associations emerge only once families accumulate at least two forms of disadvantage, while the evidence for a single marker of deprivation is comparatively fragile.

Placebo: pre-conception and *in utero* exposure included simultaneously. As a placebo test, we augment the specification by simultaneously including both *in utero* $PM_{2.5}$ and pre-conception $PM_{2.5}$ (measured over weeks 39 to 76 preceding conception[19]), each interacted with our deprivation score indicators. This design controls for the temporal autocorrelation in $PM_{2.5}$ within FSA-by-quarter cells: if the estimated associations simply reflect a persistent correlation between neighborhood pollution and unobserved determinants of health, both exposure measures should yield similarly negative coefficients. If the association is specific to *in utero* exposure, only the *in*

[18] See List et al. (2019) for interesting discussion on multiple hypothesis testing.

[19] Our results are robust to substracting gestational age from the birth date and using the 38 preceding weeks from that calculated conception date.

utero coefficients should be negative for high-deprivation families.

Results are reported in Appendix Table B.2. The pattern is consistent with the second interpretation. For Panel A (pre-conception exposure), no coefficient is negative and significant. For Panel B (*in utero* exposure), the gradient closely mirrors our main results: differentials for score = 0 and score = 1 are small and insignificant, while those for score = 2 (-3.958 grams; SE = 2.107) and score = 3 (-7.506 grams; SE = 2.681) are negative and statistically significant. The contrast between Panels A and B suggests that the deprivation gradient is specific to prenatal exposure and is not an artifact of persistent neighborhood characteristics or the temporal autocorrelation of pollution exposure.

Conley standard errors. We next examine whether the statistical inference is sensitive to spatial dependence in pollution exposure. Appendix Table B.3 reports standard errors allowing for spatial correlation (Conley, 1999) across neighborhoods within 10, 20, 30, and 40 km of one another. For both average $PM_{2.5}$ and days above 27 $\mu g/m^3$, the differential associations for score = 2 and score = 3 remain statistically significant at the 1% level across all distance cutoffs.

Alternative exposure assignment. Appendix Table B.4 re-estimates the models using monitoring stations within 20-km and 30-km buffers around each residential neighborhood centroid, rather than the 10-km buffer used in our main specification.[20] The results are remarkably stable for score = 2 and score = 3 across both exposure measures and all three buffer definitions, while estimates for scores 0 and 1 remain small and close to zero.

Alternative thresholds. Alongside our main results from Figure 3, Appendix Figure B.1 presents the estimated associations two additional threshold-based exposure measures— days above 15 and 20$\mu g/m^3$ — across deprivation score values. For families with scores of 0 and 1, associations are small and close to zero across all thresholds. The picture is strikingly different for scores

[20] Using these alternative buffers leads to an increase in sample size from ∼282,000 (10km) to ∼345,000 (20km) to ∼373,000 (30km).

of 2 and 3: the magnitude of the estimated associations increases monotonically as the threshold increases, reflecting the rarity of each threshold (98.6, 41.0, 16.3, and 5.5 days on average during the prenatal window, respectively). For families with deprivation scores of 2 and 3, the 95% confidence intervals for days above 20 μg/m^3 and 27 μg/m^3 do not overlap with those for days above 10 μg/m^3, indicating that the larger associations at higher thresholds are statistically distinguishable from those at the frequently-exceeded annual guideline. For score = 3, the differential association rises from -0.494 gram per day above 10 μg/m^3 to -2.046 grams per day above 20 μg/m^3 and -3.990 grams per day above 27 μg/m^3. These results suggest that exposure to rare but pronounced pollution episodes, rather than frequent moderate exceedances, drives the adverse associations documented for socioeconomically disadvantaged families.

Mother (sibling) fixed effects. As a final check, we re-estimate equation (1) separately for each deprivation score value, augmenting each specification with mother fixed effects. This approach identifies associations from variation in $PM_{2.5}$ exposure across pregnancies of the same mother, holding constant all time-invariant maternal characteristics. Results are reported in Appendix Table B.6. The within-mother estimates are substantially less precise, as expected given the smaller effective sample (mothers with at least two singleton births). Nevertheless, the directional pattern is consistent with our main findings: associations for scores 0 and 1 remain close to zero, while estimates for score = 2 (-9.788 grams for average $PM_{2.5}$) and score = 3 (-4.466 grams) remain negative, albeit imprecisely estimated for score = 3 given the small within-score sample of approximately 8,000 births.

4.5 High-risk births: NICU sample

Our analysis so far has focused on socioeconomic vulnerability. We next examine whether prenatal pollution exposure also has adverse effects among newborns who are biologically vulnerable. To do so, we replicate our analysis using medical records from infants admitted at birth to a Level III Neonatal Intensive Care Unit (NICU) in Quebec City. Because this sample consists primarily

of premature and medically fragile newborns, it provides a unique opportunity to assess whether prenatal pollution exposure is associated with worse birth outcomes among infants at elevated biological risk.

A potential limitation of this analysis is that admission to a NICU may itself be influenced by prenatal exposure to ambient air pollution. Recent studies report that higher exposure to $PM_{2.5}$ and other traffic-related pollutants during the weeks immediately preceding delivery is associated with an increased probability of NICU admission at birth (Seeni et al., 2019; Phiri et al., 2025). If pollution affects both neonatal health and the likelihood of admission to a NICU, conditioning on NICU admission may introduce selection bias. At the same time, the existing NICU-admission literature focuses primarily on short-term exposure immediately before birth, whereas our analysis examines the relationship between average exposure over pregnancy and two outcomes (preterm birth and low birth weight) that can reflect processes operating throughout gestation. Consistent with this distinction, we find no evidence that $PM_{2.5}$ exposure during the week preceding delivery is associated with either preterm birth or low birth weight in our data. Although this finding cannot rule out selection through other neonatal conditions that influence NICU admission, it suggests that selection bias arising through the birth outcomes considered here is likely to be limited. Consequently, the extent to which selection into the NICU sample is driven by the cumulative exposure measures used in this study remains uncertain. The estimates reported below should therefore be interpreted as associations among newborns admitted to a NICU rather than as causal effects for the population of all births. To our knowledge, previous studies have examined whether prenatal pollution exposure increases the probability of NICU admission, but have not investigated whether pollution exposure is associated with birth outcomes conditional on NICU admission.

Detailed information on the data is provided by Beltempo et al. (2022, 2023). The sample consists of 3,812 newborns admitted at birth to a Level III NICU in Québec city between April 2008 and March 2013.[21] Table C.1 presents the main characteristics of the sample. Relative to newborns in maternity wards, NICU patients were more likely to be male (57% *vs* 51%), had a

[21] Level III NICUs provide intensive care by pediatricians-neonatalogists available 24 hours per day, with dedicated nursing staff, and immediate access to all medical specialties.

substantially lower mean birth weight (2,983 grams *vs* 3,348 grams), and were born at an earlier gestational age (37.2 *vs* 38.9 weeks). Consistent with these differences, 26% were classified as low birth weight and 31% were born preterm, compared with only 6% and 7%, respectively, among newborns admitted to maternity wards.

Interestingly, newborns in the NICU sample experienced lower $PM_{2.5}$ exposure than newborns in the province-wide maternity sample. Their average $PM_{2.5}$ exposure was 8.10 $\mu g/m^3$ (versus 9.72 $\mu g/m^3$), they experienced fewer days with $PM_{2.5}$ concentrations above 10 $\mu g/m^3$ (68.6 versus 98.6), and fewer days with concentrations above 27 $\mu g/m^3$ (2.95 versus 5.51). These differences may reflect, at least in part, the geographic composition of the NICU sample, which is mostly drawn from the Québec City region rather than the entire province, together with differences in the timing and duration of pregnancy associated with the shorter gestational age of NICU newborns.

Two-Way Fixed Effects Estimates. Appendix Table C.2 reports the estimated effects of prenatal $PM_{2.5}$ exposure obtained from the same two-way fixed-effects specification used in the province-wide analysis, estimated here on the NICU sample.[22] Holding the empirical specification constant enables a more direct comparison between the two populations, allowing differences in the estimated effects to be interpreted as being consistent with greater biological vulnerability rather than changes in model specification. Consistent with our cumulative vulnerability hypothesis, prenatal $PM_{2.5}$ exposure appears to have more pronounced effects among biologically vulnerable newborns than in the population as a whole. A one-unit increase in average $PM_{2.5}$ exposure is associated with a 32.9-gram reduction in birth weight, corresponding to approximately 1.1% of the mean birth weight. The same increase is associated with a 2.3 percentage-point increase in the probability of low birth weight and a 1.8 percentage-point increase in the probability of preterm birth, representing increases of approximately 8.8% and 5.2%, respectively, relative to sample means. We also find suggestive evidence that *in utero* $PM_{2.5}$ exposure is associated with lower gestational age and with a higher probability of having an Apgar score below 7, although these estimates are

[22] The controls included slightly differ from those included in the specification estimated in the main sample based on data availability. See details in Appendix Table C.2. We also could not construct our deprivation index in the NICU sample given the information available.

imprecise.

An additional insight from these estimates is that adverse associations emerge not only for days exceeding the 27 μg/m^3 daily standard, but also for days above the 10 μg/m^3 annual guideline.[23] Whereas the population-based analysis suggests that the largest effects are concentrated during relatively rare high-pollution episodes, the NICU analysis indicates that medically fragile newborns may be adversely affected by more moderate increases in pollution exposure. For example, each additional day with $PM_{2.5}$ concentrations above 10 μg/m^3 is associated with a 2.03-gram reduction in birth weight and a 0.2 percentage-point increase in the probability of preterm birth. Stronger associations are estimated for days with concentrations above the daily threshold of 27 μg/m^3, these estimates are also substantially larger than those from our estimations in the general population. Although the NICU sample is selective and the estimates should be interpreted with caution, they reinforce our broader conclusion that the health consequences of prenatal $PM_{2.5}$ exposure are disproportionately borne by the most vulnerable newborns, whether their vulnerability is socioeconomic or biological.

Joint Modeling of Preterm and Low Birth Weight. Although low birth weight and preterm birth are positively correlated, they capture distinct dimensions of neonatal health. Preterm birth reflects shortened gestation, whereas low birth weight may result either from prematurity or from impaired fetal growth. Consequently, some infants are born preterm without being classified as low birth weight, while others are born at term but nevertheless weigh less than 2,500 grams. Because *in utero* $PM_{2.5}$ exposure may affect gestational duration, fetal growth, or both, analyzing these outcomes jointly provides a more complete assessment of its health consequences than modeling each outcome separately.

Table 5 illustrates this distinction. Among the 952 infants classified as low birth weight, 664 (69.7%) were also born preterm, while 288 (30.3%) reached full term despite weighing less than 2,500 grams, suggesting fetal growth restriction rather than prematurity. Conversely, among the 784 preterm infants, 664 (84.7%) were also classified as low birth weight, whereas 120 (15.3%)

[23] Results from each regression are reported on a separate line in Table C.2.

exceeded the clinical threshold for low birth weight despite shortened gestation. Thus, although the two outcomes overlap substantially, neither is a perfect proxy for the other. Their imperfect concordance provides empirical support for modeling low birth weight and preterm birth jointly rather than separately.

Table 5: Low birth weight and preterm birth in the Quebec NICU

	Preterm birth		
Low birth weight	No	Yes	Total
No	2,172	120	2,292
Yes	288	664	952
Total	2,460	784	3,244

Notes: The sample used in the spatial probit model differs slightly from that used in the two-way fixed-effects model because observations located in FSAs without contiguous neighbors cannot be included in the spatial adjacency matrix.

Bivariate Probit Model Estimation. To facilitate comparisons, we use the same set of covariates and fixed effects as in the two-way fixed-effects models estimated in our NICU sample, and model low birth weight and preterm birth jointly using a spatial bivariate probit specification. Modeling the two outcomes simultaneously allows their latent propensities to be correlated, thereby accounting for unobserved maternal and fetal characteristics that jointly influence both birth outcomes. An additional advantage of the joint specification is that it permits direct estimation of the effect of prenatal $PM_{2.5}$ exposure on the probability of simultaneously experiencing low birth weight and preterm birth, a quantity that cannot be obtained from separate univariate models. Our NICU sample is particularly well-suited for this exercise, given that these two outcomes are not as rare as in the general sample of births. Furthermore, we find that prematurity is more sensitive to *in utero* exposure to $PM_{2.5}$ than the general population.

To account for geographic dependence, the model incorporates area-specific spatial random effects through an intrinsic bivariate conditional autoregressive (BICAR) prior.[24] This specification smooths latent risks across neighboring forward sortation areas (FSAs) while allowing the spatial

[24] We follow the model proposed by Neelon et al. (2014). Details are provided in Appendix C.2.

effects for low birth weight and preterm birth to be correlated. Consequently, the model captures two distinct sources of dependence: the correlation between the two birth outcomes within an infant and the spatial correlation among neighboring geographic areas. Holding the empirical specification otherwise constant facilitates a direct comparison with the province-wide results, so that differences in the estimated effects primarily reflect differences in the study populations rather than changes in the econometric framework. The complete statistical specification and estimation procedure are presented in Appendix C.2.

It is useful to assess the extent of dependence between the two birth outcomes before examining the estimated effects *in utero* $PM_{2.5}$ has on them. Table C.3 reports the posterior summaries of the key model parameters. The estimated latent correlation between low birth weight and preterm birth is remarkably high ($\rho = 0.888$; 95% credible interval: 0.867–0.907), indicating that the two outcomes remain strongly associated even after accounting for observed covariates and spatial random effects and providing empirical support for jointly modeling them. In contrast, the posterior distribution of the covariance between the two spatial random effects is centered near zero and is estimated imprecisely, suggesting limited evidence of additional cross-outcome spatial dependence beyond that captured by the latent correlation parameter.

Figure C.1 displays the posterior distributions of the $PM_{2.5}$ coefficients for each exposure measure and birth outcome. Rather than relying solely on posterior means and 95% credible intervals, these distributions quantify the posterior probability that prenatal $PM_{2.5}$ exposure increases the risk of adverse birth outcomes. For average $PM_{2.5}$ exposure, the posterior probability of a positive effect is 97.1% for low birth weight and 84.0% for preterm birth. Similar evidence is obtained for the number of days with $PM_{2.5}$ concentrations exceeding 10 $\mu g/m^3$, with posterior probabilities of 97.3% and 81.0%, respectively. In contrast, days exceeding 27 $\mu g/m^3$ provide comparatively weak evidence of an association with low birth weight (61.2%) but strong evidence of an association with preterm birth (96.7%). Taken together, the posterior distributions suggest that different dimensions of prenatal $PM_{2.5}$ exposure may influence neonatal health through distinct biological mechanisms for this sample of particularly vulnerable infants. Measures reflecting sustained exposure, namely

average $PM_{2.5}$ concentration and the number of days exceeding the annual guideline of 10 $\mu g/m^3$, exhibit the strongest evidence of an association with low birth weight, whereas episodic exposure to high pollution levels (days above 27 $\mu g/m^3$) appears to be more closely associated with preterm birth. Although these patterns should be interpreted cautiously, they are consistent with the hypothesis that chronic and acute pollution exposures affect different aspects of fetal development, especially for babies that are particularly at-risk for various biological reasons.

Table C.4 reports the corresponding posterior distributions of the average marginal effects for average $PM_{2.5}$ exposure. A one-unit increase in average $PM_{2.5}$ is associated with a 2.2 percentage-point increase in the probability of low birth weight and a 1.3 percentage-point increase in the probability of preterm birth for NICU babies. Importantly, the spatial bivariate probit model also permits direct estimation of the joint probability of simultaneously experiencing both adverse birth outcomes. The posterior mean indicates that a one-unit increase in average $PM_{2.5}$ exposure increases the probability of being both preterm and low birth weight by approximately 1.5 percentage points. This joint effect is a distinctive feature of the bivariate specification and cannot be recovered from separate univariate models.

As a robustness check, Table C.5 compares the average marginal effects obtained from separate univariate logit models with those from the spatial bivariate probit model, using the same sample of NICU newborns. Overall, the estimated marginal effects are remarkably similar across specifications, indicating that the substantive conclusions are not sensitive to the choice of econometric model. For average $PM_{2.5}$ exposure, both models imply economically meaningful increases in the probabilities of low birth weight and preterm birth, with the spatial bivariate probit yielding only slightly smaller marginal effects. Likewise, both approaches suggest that sustained exposure, as measured by average $PM_{2.5}$ and the number of days exceeding 10 $\mu g/m^3$, is more strongly associated with low birth weight, whereas high-pollution episodes (days above 27 $\mu g/m^3$) exhibit a stronger association with preterm birth.

The close agreement between the two approaches suggests that accounting for outcome dependence and spatial correlation does not materially alter the estimated marginal effects of prenatal

$PM_{2.5}$ exposure.[25] Rather, the principal contribution of the spatial bivariate probit model lies in its richer inferential framework. By jointly modeling low birth weight and preterm birth, the model accounts for their substantial latent correlation, accommodates residual spatial dependence across neighboring FSAs, and permits direct estimation of the effect of prenatal $PM_{2.5}$ exposure on the joint probability of experiencing both adverse birth outcomes.

The estimated correlation between the latent propensities for low birth weight and preterm birth is high ($\rho = 0.888$), indicating that the two outcomes share substantial unobserved determinants after controlling for observed maternal, infant, and neighborhood characteristics. By contrast, the posterior distribution of the covariance between the outcome-specific spatial random effects is centered near zero, providing little evidence that the remaining spatial heterogeneity is jointly shared across the two outcomes. Together, these results suggest that the dependence between low birth weight and preterm birth arises primarily from shared individual-level risk factors rather than common unobserved neighborhood characteristics.

5 Conclusion

Our findings suggest that the health consequences of prenatal $PM_{2.5}$ exposure in Quebec differ markedly from those documented in many higher-pollution settings. At the population level, we find little evidence that prenatal $PM_{2.5}$ exposure substantially affects birth outcomes on average. Estimated effects on birth weight and gestational age are modest, and we find essentially no association with low birth weight, preterm birth, or small-for-gestational-age status. These results are consistent with a growing body of evidence reporting limited average effects of prenatal $PM_{2.5}$ exposure in relatively clean environments. For example, Jahanshahi et al. (2022) find no statistically significant effects in Northern Ireland, while Ling (2023) report similarly modest average effects in Norway adopting a comparable empirical strategy. Our findings reinforce the view that average

[25] The results presented in Table C.5 are also quite close to those from columns 2 and 5 of Table C.2, which includes births in FSAs that do not have contiguous neighbours, and which are removed from the sample for the estimation of the spatial probit model.

effects of prenatal $PM_{2.5}$ exposure may be small in low-pollution settings. An important direction for future work may be to test whether our estimates represent lower bounds in our specific context using an instrumental variable framework or quasi-experimental design based on policy-induced variation in exposure, given that our empirical approach based on monitoring station readings and the inclusion of month-of-birth, year-of-birth and quarter-of-birth×neighbourhood fixed effects may still suffer from attenuation bias due to classical measurement error (e.g., Arceo et al., 2016). Importantly, however, our results demonstrate that null average effects should not be interpreted as evidence that pollution does not matter in such environments. Rather, negligible population-level associations can conceal substantial heterogeneity in vulnerability to *in utero* pollution exposure across different groups.

Indeed, our cumulative vulnerability analysis reveals a clear socioeconomic gradient in the impact of *in utero* pollution exposure, beyond what exposure levels could explain. Adverse associations are concentrated among children born into disadvantaged families, particularly those combining multiple forms of deprivation. Estimated effects are small and statistically insignificant for families with no or only one marker of deprivation, but increase markedly for families with two and three markers. This pattern suggests that vulnerability to environmental risks is multidimensional and that socioeconomic disadvantages combine to amplify susceptibility to prenatal pollution. The estimated effect for the most disadvantaged families in our sample (6.64 grams per $\mu g/m^3$) is comparable to, and in some cases exceeds, population-average estimates reported in substantially more polluted settings, despite Quebec's comparatively low ambient $PM_{2.5}$ concentrations. Li and Zhang (2024) report an effect of 6.42 grams per additional $\mu g/m^3$ of $PM_{2.5}$ in rural China, where average exposure levels exceed 24 $\mu g/m^3$. Basu et al. (2004) estimate effects ranging from 4.04 to 4.35 grams per unit in California, where average exposure levels are between 14.5 and 18.2 $\mu g/m^3$ — 49% to 87% above our sample mean. These comparisons should nevertheless be interpreted cautiously because differences in pollution mixtures, empirical specifications, and exposure distributions limit direct cross-study comparisons.

Our analysis focused on newborns in a neonatal intensive care unit then demonstrates that

vulnerability extends beyond socioeconomic disadvantage. Applying the same empirical specification to this population of biologically vulnerable newborns yields substantially larger estimated effects than those observed in the province-wide sample. Among infants admitted to intensive care, prenatal $PM_{2.5}$ exposure is associated with meaningful reductions in birth weight and higher probabilities of low birth weight and preterm birth. Moreover, adverse associations emerge even for days with $PM_{2.5}$ concentrations exceeding 10 $\mu g/m^3$ (the threshold for average annual exposure in Canada for most of our sample years), whereas comparable exposure levels have relatively modest effects in the population as a whole. Together, the cumulative vulnerability and NICU analyses indicate that the health consequences of prenatal pollution depend not only on ambient exposure levels but also on the susceptibility of the exposed population. Vulnerability therefore appears to operate through multiple dimensions, including both socioeconomic disadvantage and underlying biological fragility. Moreover, our spatial bivariate analysis in this sample further strengthens these conclusions. Although the estimated marginal effects are very similar to those obtained from separate models, jointly modeling low birth weight and preterm birth reveals a strong residual dependence between the two outcomes even after accounting for observed characteristics and spatial effects. The joint specification also permits direct estimation of the probability that prenatal $PM_{2.5}$ exposure simultaneously increases the risks of both preterm birth and low birth weight, providing information that cannot be recovered from separate univariate models. These results also suggest that chronic and acute pollution exposures affect distinct aspects of fetal development.

These findings have important policy implications. While average ambient $PM_{2.5}$ concentrations in Quebec remain low by international standards, increasingly frequent wildfire smoke episodes may generate short-lived but substantial increases in exposure that disproportionately affect vulnerable populations. At the same time, the socioeconomic characteristics defining our cumulative vulnerability index closely resemble the eligibility criteria of prenatal support programs such as OLO (*Oeuf-Lait-Orange*, providing nutritional support and advice to low-income families that are expecting or have a newborn). Previous evaluations have shown that these programs im-

prove birth outcomes.[26] Whether they also mitigate the adverse consequences of prenatal pollution exposure—or whether pollution diminishes part of their benefits—remains an important question for future research.

Finally, our results contribute to a broader literature demonstrating that the consequences of early-life pollution exposure may extend well beyond birth. Studies such as Klauber et al. (2024), Isen et al. (2017), Gehrsitz (2017), and Pestel and Wozny (2021) suggest that even modest improvements in air quality can generate lasting gains in health and human capital. In this context, the heterogeneous associations documented here may represent only part of the long-run burden associated with prenatal pollution exposure. Overall, our findings suggest that environmental inequalities persist even in relatively clean environments. The remaining burden of prenatal $PM_{2.5}$ exposure falls disproportionately on children facing multiple forms of socioeconomic disadvantage or biological vulnerability, highlighting the importance of targeting preventive interventions toward those who are most susceptible rather than toward the average child.

[26] See Haeck and Lefebvre (2016) for evidence.

References

Alexander, D. and Schwandt, H. (2022). The impact of car pollution on infant and child health: Evidence from emissions cheating. *The Review of Economic Studies*, 89(6):2872–2910.

Almond, D., Chay, K. Y., and Lee, D. S. (2005). The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083.

Arceo, E., Hanna, R., and Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. *The Economic Journal*, 126(591).

Basu, R., Woodruff, T. J., Parker, J. D., Saulnier, L., and Schoendorf, K. C. (2004). Comparing exposure metrics in the relationship between pm2. 5 and birth weight in california. *Journal of Exposure Science & Environmental Epidemiology*, 14(5):391–396.

Beltempo, M., Bresson, G., Étienne, J.-M., and Lacroix, G. (2022). Infections, accidents and nursing overtime in a neonatal intensive care unit. *The European Journal of Health Economics*, 23(4):627–643.

Beltempo, M., Bresson, G., and Lacroix, G. (2023). Using machine learning to predict nosocomial infections and medical accidents in a nicu. *Health and Technology*, 13(1):75–87.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Bharadwaj, P., Gibson, M., Zivin, J. G., and Neilson, C. (2017). Gray matters: Fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists*, 4(2):505–542.

Bharadwaj, P., Lundborg, P., and Rooth, D.-O. (2018). Birth weight in the long run. *Journal of Human Resources*, 53(1):189–231.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439.

Brook, R. D., Rajagopalan, S., Pope III, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., et al. (2010). Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the american heart association. *Circulation*, 121(21):2331–2378.

Chay, K. and Greenstone, M. (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The quarterly journal of economics*, 118(3):1121–1167.

Chen, L., Chen, Y., and Lin, J. (2026). What doesn't kill you makes you shorter: The effect of prenatal air pollution on height in adulthood. *Journal of Economic Behavior & Organization*, 248:107633.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Currie, J., Neidell, M., and Schmieder, J. F. (2009). Air pollution and infant health: Lessons from new jersey. *Journal of health economics*, 28(3):688–703.

Currie, J. and Walker, R. (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics*, 3(1):65–90.

Cyrys, J., Pitz, M., Bischof, W., Wichmann, H.-E., Heinrich, J., et al. (2004). Relationship between indoor and outdoor levels of fine particle mass, particle number concentrations and black smoke under different ventilation conditions. *Journal of Exposure Science & Environmental Epidemiology*, 14(4):275–283.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12).

Gehrsitz, M. (2017). The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management*, 83:121–144.

Ghafarianghadim, M. and Heutel, G. (2026). Air pollution and early childhood outcomes: Evidence from georgia childcare providers. Technical report, National Bureau of Economic Research.

Glinianaia, S. V., Rankin, J., Bell, R., Pless-Mulloli, T., and Howel, D. (2004). Particulate air pollution and fetal health: a systematic review of the epidemiologic evidence. *Epidemiology*, 15(1):36–45.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of economic literature*, 51(3):689–730.

Graff Zivin, J., Neidell, M., Sanders, N. J., and Singer, G. (2023). When externalities collide: Influenza and pollution. *American Economic Journal: Applied Economics*, 15(2).

Haeck, C. and Lefebvre, P. (2016). A simple recipe: The effect of a prenatal nutrition program on child health at birth. *Labour Economics*, 41:77–89.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pages 65–70.

Isen, A., Rossin-Slater, M., and Walker, W. R. (2017). Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy*, 125(3):848–902.

Jahanshahi, B., Johnston, B., McVicar, D., McGovern, M., O'Reilly, D., Rowland, N., and Vlachos, S. (2022). Prenatal exposure to pm2. 5 and infant birth outcomes: Evidence from a population-wide database. Working paper, Queen's University Belfast Law Research Paper, Forthcoming.

Jans, J., Johansson, P., and Nilsson, J. P. (2018). Economic status, air quality, and child health: Evidence from inversion episodes. *Journal of Health Economics*, 61:220–232.

Jedrychowski, W., Majewska, R., Spengler, J., Camann, D., Roen, E., and Perera, F. (2017). Prenatal exposure to fine particles and polycyclic aromatic hydrocarbons and birth outcomes: a two-pollutant approach. *International archives of occupational and environmental health*, 90:255–264.

Klauber, H., Holub, F., Koch, N., Pestel, N., Ritter, N., and Rohlf, A. (2024). Killing prescriptions softly: Low emission zones and child health from birth to school. *American Economic Journal: Economic Policy*, 16(2):220–248.

Kraemer, S. (2000). The fragile male. *Bmj*, 321(7276):1609–1612.

Kramer, M. S. (1987). Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the world health organization*, 65(5):663.

Li, L. and Zhang, X. (2024). The causal impact of fetal exposure to pm2. 5 on birth outcomes: Evidence from rural china. *Economics & Human Biology*, 53:101380.

Ling, X. (2023). The effect of ambient air pollution on birth outcomes in norway. *BMC Public Health*, 23(1):2248.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.

Neelon, B., Anthopolos, R., and Miranda, M. L. (2014). A spatial bivariate probit model for correlated binary data with application to adverse birth outcomes. *Statistical Methods in Medical Research*, 23(2):119–133.

Palma, A., Petrunyk, I., and Vuri, D. (2022). Prenatal air pollution exposure and neonatal health. *Health Economics*, 31(5):729–759.

Pestel, N. and Wozny, F. (2021). Health effects of low emission zones: evidence from german hospitals. *Journal of Environmental Economics and Management*, 109:102512.

Phiri, Y. V. A., Canty, T., Nobles, C., Ring, A. M., Nie, J., and Mendola, P. (2025). Neonatal intensive care admissions and exposure to satellite-derived air pollutants in the United States, 2018. *Scientific Reports*, 15:420.

Sanders, N. J. and Stoecker, C. (2015). Where have all the young men gone? using sex ratios to measure fetal death rates. *Journal of health economics*, 41:30–45.

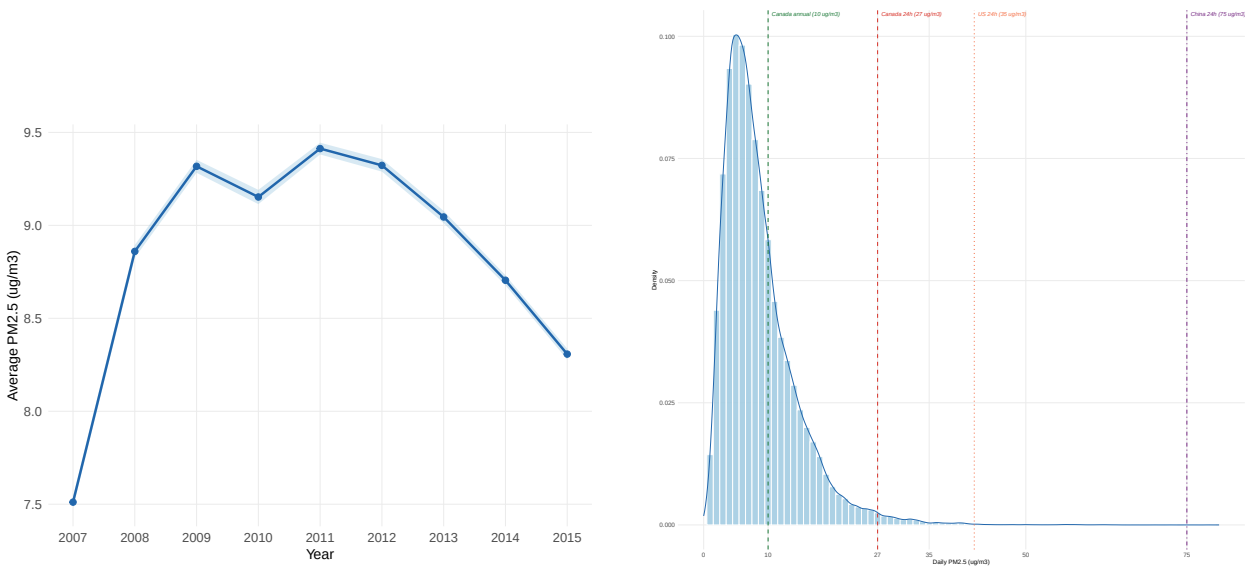
Seeni, I., Williams, A., Nobles, C., Chen, Z., Sherman, S., and Mendola, P. (2019). Acute air pollution exposure and NICU admission: A case-crossover analysis. *Annals of Epidemiology*, 37:64–70.e2.

Simeonova, E., Currie, J., Nilsson, P., and Walker, R. (2021). Congestion pricing, air pollution, and children's health. *Journal of Human Resources*, 56(4):971–996.

van den Hooven, E. H., de Kluizenaar, Y., Pierik, F. H., Hofman, A., van Ratingen, S. W., Zandveld, P. Y., Mackenbach, J. P., Steegers, E. A., Miedema, H. M., and Jaddoe, V. W. (2011). Air pollution, blood pressure, and the risk of hypertensive complications during pregnancy: the generation r study. *Hypertension*, 57(3):406–412.

You, S. and Chou, S.-Y. (2025). Air pollution and health of working-age population: Evidence from thermal inversion. *Economics & Human Biology*, 59:101558.

A Air pollution trend



(a) Annual average $PM_{2.5}$

(b) Distribution of daily $PM_{2.5}$

Figure A.1: $PM_{2.5}$ trend and distribution, 2007–2015

Notes: Panel (a) shows the province-wide annual average of daily $PM_{2.5}$ concentrations across all monitoring stations, with 95% confidence bands. Panel (b) shows the distribution of daily station-level $PM_{2.5}$ readings, with vertical lines indicating the Canadian annual standard ($10\,\mu g/m^3$), the Canadian 24-hour standard ($27\,\mu g/m^3$), the US 24-hour standard ($35\,\mu g/m^3$), and the Chinese 24-hour standard ($75\,\mu g/m^3$).

B Robustness Analysis

Table B.1: Birth weight (grams) by deprivation score

Score	Avg. $PM_{2.5}$	Days > 27 $\mu g/m^3$
0 (least deprived)	-0.283	-0.289
	(0.947)	(0.559)
1	0.454	1.394***
	(0.297)	(0.430)
2	-2.844***	-1.220**
	(0.410)	(0.547)
3 (most deprived)	-6.359***	-3.990***
	(0.643)	(0.741)

Notes: Each row reports the differential effect relative to score=0, estimated from a single regression interacting $PM_{2.5}$ exposure with indicators for score values 1, 2, and 3 (score=0 as reference). Each coefficient $\hat{\beta}_s$ captures the effect for score s relative to score=0. Deprivation score = family income below $45,000 + low-income neighborhood + no post-secondary education. All specifications include demographic controls, weather controls, neighborhood-level controls, month FE, year FE, and neighborhood-by-quarter FE. SE clustered at neighborhood level. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.2: Placebo test: pre-conception vs. *in utero* exposure (birth weight, grams)

	Deprivation score			
	0 (least)	1	2	3 (most)
Panel A. Pre-conception exposure ($\hat{\beta}_s^{before}$)				
Avg. $PM_{2.5}$	-1.829	3.225*	1.177	1.261
	(1.763)	(1.927)	(2.107)	(2.498)
Days > 27 $\mu g/m^3$	-0.431	1.018*	-0.193	-1.422*
	(0.563)	(0.544)	(0.608)	(0.813)
Panel B. In utero exposure ($\hat{\beta}_s^{during}$)				
Avg. $PM_{2.5}$	1.893	-2.632	-3.958*	-7.506***
	(1.811)	(1.956)	(2.107)	(2.681)
Days > 27 $\mu g/m^3$	-0.415	0.738	-1.118*	-2.983***
	(1.094)	(0.476)	(0.591)	(0.804)

Notes: Both exposure measures are included simultaneously in a single regression interacted with indicators for deprivation score values 0–3 (score=0 as reference). Pre-conception exposure is measured over the 38 weeks preceding conception; in utero exposure is measured over the 38 weeks preceding birth. Coefficients reported are differential associations relative to score=0 for each exposure measure. All specifications include the full set of demographic, weather, and neighborhood controls, month FE, year FE, and neighborhood-by-quarter FE. SE clustered at the neighborhood level. $*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.3: Conley spatial standard errors (Avg. $PM_{2.5}$, birth weight)

		Conley SE at cutoff			
Score	Estimates	10 km	20 km	30 km	40 km
0	-0.283	(1.121)	(1.192)	(1.146)	(1.231)
1	0.454	(0.282)	(0.278)	(0.379)	(0.209)**
2	-2.844	(0.523)***	(0.527)***	(0.537)***	(0.386)***
3	-6.359	(0.847)***	(0.703)***	(0.810)***	(0.527)***

Notes: Differential effects relative to score=0 ($\hat{\beta}_s$) from Table B.1, controlling for the same set of variables and fixed effects as in Table B.1. Each SE column reports Conley (1999) spatial HAC standard errors allowing for spatial correlation across neighborhoods within the indicated distance cutoff. Stars indicate significance of the point estimate using the corresponding Conley SE. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.4: Robustness to alternative exposure assignment (birth weight, grams)

Score	Baseline (10 km)	20 km	30 km
Panel A. Average $PM_{2.5}$			
0	-0.283	0.483	0.572
	(0.947)	(0.644)	(1.116)
1	0.454	1.018**	-0.026
	(0.297)	(0.484)	(0.297)
2	-2.844***	-2.976***	-4.101***
	(0.410)	(0.678)	(0.438)
3	-6.359***	-6.050***	-7.818***
	(0.643)	(0.895)	(0.649)
Panel B. Days > 27 $\mu g/m^3$			
0	-0.289	0.189	0.133
	(0.559)	(0.602)	(0.978)
1	1.394***	0.907*	-0.012
	(0.430)	(0.487)	(0.308)
2	-1.220**	-2.705***	-3.962***
	(0.547)	(0.667)	(0.446)
3	-3.990***	-5.724***	-7.645***
	(0.741)	(0.857)	(0.649)
N	282,000	345,000	373,000

Notes: Each cell reports the differential effect relative to score=0 ($\hat{\beta}_s$) from a single regression with score-by-exposure interactions (score=0 as reference). Columns differ in the buffer distance used to assign $PM_{2.5}$ exposure. All specifications include the full set of controls and fixed effects from Table B.1. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

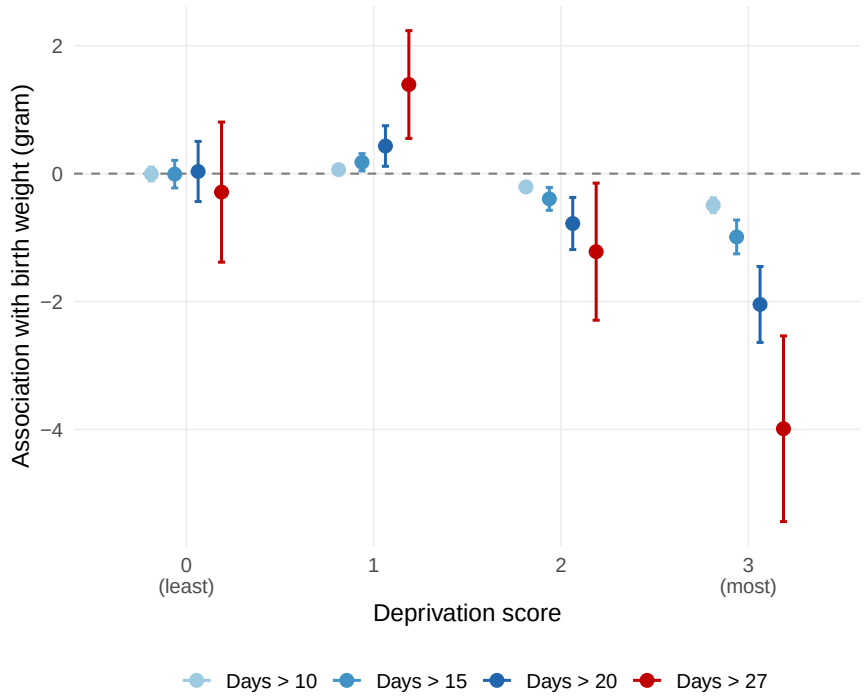


Figure B.1: Associations between threshold-based $PM_{2.5}$ exposure and birth weight by deprivation score

Notes: Notes: Each point reports a differential association relative to score = 0 (except score = 0 which reports the baseline). Bars show 95% CI. SE clustered at FSA level. Deprivation score = family income < CAD$45,000 +low-income neighborhood + no post-secondary (0-3).

Table B.5: Multiple hypothesis testing corrections

	p-value		
Score	Unadjusted	Bonferroni	FDR
Panel A. Average $PM_{2.5}$			
1	0.11	1.00	0.11
2	< 0.001	< 0.001	< 0.001
3	< 0.001	< 0.001	< 0.001
Panel B. Days $> 10\,\mu g/m^3$			
1	0.028	0.253	0.032
2	< 0.001	< 0.001	< 0.001
3	< 0.001	< 0.001	< 0.001
Panel C. Days $> 27\,\mu g/m^3$			
1	< 0.001	0.013	< 0.001
2	0.021	0.192	0.027
3	< 0.001	< 0.001	< 0.001

Notes: p-values for the differential association between $PM_{2.5}$ exposure and birth weight for each deprivation score value relative to score=0 (reference). The family of hypotheses includes scores 1, 2, and 3 for each of the three exposure measures — Average $PM_{2.5}$, Days $> 10\,\mu g/m^3$, and Days $> 27\,\mu g/m^3$ (nine hypotheses in total), corrected jointly. Bonferroni correction (Holm, 1979) multiplies each p-value by the number of hypotheses tested (capped at 1). FDR applies the Benjamini and Hochberg (1995) procedure. Unadjusted p-values correspond to the estimates in Table B.1.

Table B.6: Mother fixed effects estimates (birth weight, grams)

Score	Avg. $PM_{2.5}$	Days $> 27\,\mu g/m^3$	N
0 (least deprived)	-1.992	1.474	20,000
	(4.341)	(2.055)	
1	2.737	1.103	50,000
	(2.524)	(1.070)	
2	-9.788*	-2.148	18,000
	(5.872)	(2.337)	
3 (most deprived)	-4.466	-2.538	8,000
	(9.479)	(3.167)	

Notes: Each row reports coefficients from a separate regression estimated on the subsample with the indicated deprivation score value, including mother fixed effects and restricted to mothers with at least two singleton births in the sample. The fixed effects included in Table B.1 are also included. Standard errors are clustered at the mother level. N reports the number of births in each score subsample. Within-mother estimation absorbs time-invariant maternal characteristics but substantially reduces statistical power; none of the estimates are significant at conventional levels except score=2 for Avg. $PM_{2.5}$ ($p<0.10$). The sign and ordering across score values are directionally consistent with Table B.1 for scores 2 and 3. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

C Analyse using NICU Data

C.1 Two-way fixed effets models: NICU data

Table C.1: Descriptive Statistics: NICU Sample

	Mean	St. Dev.	Min	Max
Panel A. Birth outcomes				
Gestational age (weeks)	37.22	3.20	26	42
Birth weight (grams)	2,983.88	806.32	350	5,200
Low birth weight (<2,500 g)	0.26	0.44	0	1
Very low birth weight (<1,500 g)	0.06	0.23	0	1
Apgar score < 7	0.10	0.30	0	1
Preterm birth	0.31	0.46	0	1
Male infant	0.57	0.49	0	1
Panel B. Neighborhood characteristics				
Population size (FSA)	21,585.33	15,628.28	2,711	83,586
Average property value	210,640.40	81,337.55	88,368	646,855
Median household income	59,899.91	17,527.17	25,737	121,856
Immigrant share	0.04	0.03	0.005	0.14
Employment rate	0.642	0.071	0.421	0.786
Panel C. Pollution exposure				
Average $PM_{2.5}$	8.19	1.78	2.41	15.84
# Days above 10 $\mu g/m^3$	68.60	27.25	4	187
# Days above 27 $\mu g/m^3$	2.95	2.18	0	25

Table C.2: Parameter estimates, NICU sample
In utero exposure and birth outcomes

Dependent Variables:	Birth weight (1)	LBW (2)	VLBW (3)	GA (4)	Preterm (5)	Apgar<7 (6)
Variables						
Average $PM_{2.5}$	-32.93*	0.023***	0.004	-0.097	0.0182*	0.0126
	(16.71)	(0.008)	(0.005)	(0.073)	(0.010)	(0.008)
# Days > annual threshold	-2.025*	0.002***	0.0002	-0.005	0.001	0.001
	(1.167)	(0.001)	(0.0003)	(0.005)	(0.001)	(0.0005)
# Days > daily threshold	-17.90**	0.006	-0.001	-0.049*	0.010**	0.004
	(8.475)	(0.004)	(0.003)	(0.026)	(0.005)	(0.004)
Mean Dependent	2,983.8	0.26	0.06	37.22	0.35	0.099
Observations			3,812			

Notes: Controls include child sex, delivery mode, DRG severity index, weather and zip code characteristics. All specifications include FSA×quarter, month, and year fixed effects. Standard errors are clustered at FSA level level.Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. p<0.01.

Table C.3: Posterior summaries of $PM_{2.5}$ key model parameters

	Posterior mean	2.5%	97.5%
Preterm	0.042	-0.037	0.122
Low birth weight	0.080	-0.007	0.160
$\rho = corr(y^*_{1ij}, y^*_{2ij} \mid \phi_i)$	0.888	0.867	0.907
$\Sigma_{12} = Cov(\phi_{1i}, \phi_{2i})$	-0.124	-1.895	0.767

Table C.4: Posterior distribution of average marginal effects of prenatal $PM_{2.5}$ exposure

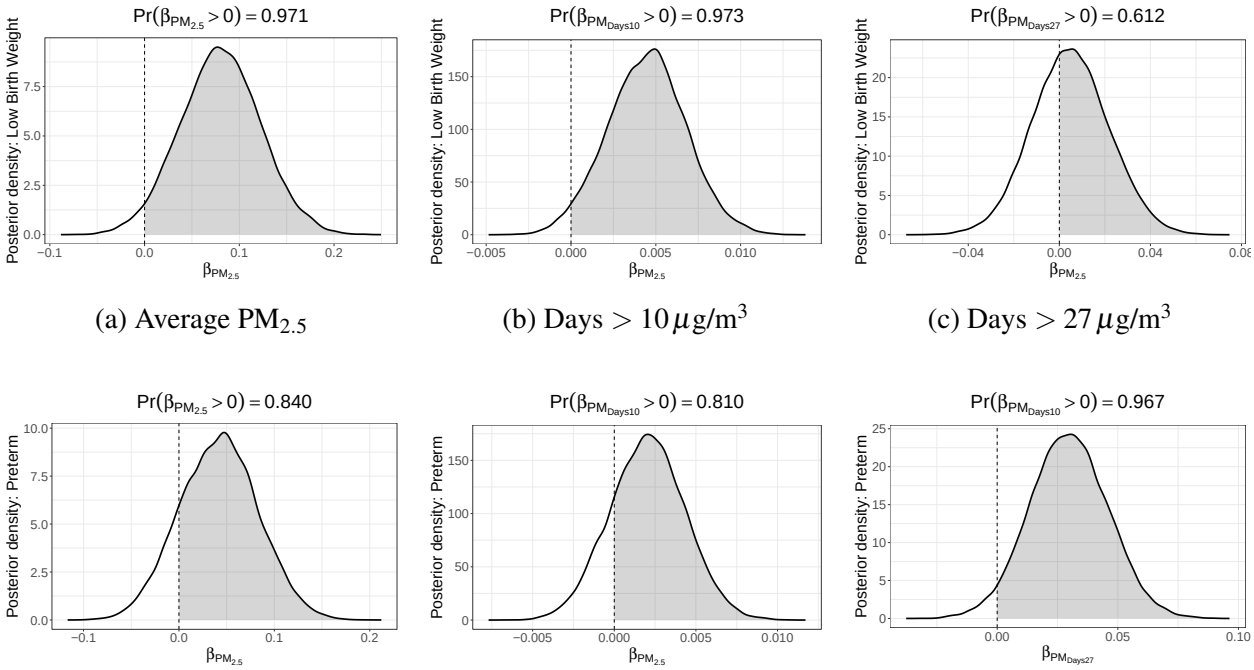
	Posterior mean	2.5%	25%	50%	75%	97.5%
Prob(PT=1)	0.013	-0.011	0.004	0.013	0.021	0.038
Prob(LBW=1)	0.022	-0.001	0.014	0.022	0.030	0.046
Prob(PT=1 and LBW=1)	0.015	-0.002	0.009	0.015	0.022	0.034

Notes: Controls include child sex, delivery mode, DRG severity index, weather and zip code characteristics. All specifications include FSA×quarter, month, and year fixed effects.

Table C.5: Robustness of estimated marginal effects across alternative model specifications

	Logit Model		Spatial Probit Model	
Dependent variable:	LBW	Preterm	LBW	Preterm
Model:	(1)	(2)	(3)	
Average Marginal Effects				
Average $PM_{2.5}$	0.027**	0.015	0.022**	0.013*
	(0.012)	(0.014)	(0.012)	(0.012)
Days above 10 $\mu g/m^3$	0.001**	0.001	0.001	0.0005
	(7e-04)	(8e-04)	(6.4e-04)	(8.1e-4)
Days above 27 $\mu g/m^3$	0.003	0.010*	0.001	0.008**
	(0.005)	(0.005)	(0.004)	(0.004)
Mean dep. var.	0.242	0.293	0.242	0.293
Observations	3,244		3,244	

Notes: Average marginal effects from logit models estimated using `feglm` (`fixest`). Standard errors are computed using the delta method with a variance–covariance matrix clustered at the FSA level. Controls include child sex, delivery mode, DRG severity index, weather conditions, and neighborhood characteristics. All three specifications include FSA×quarter, month, and year fixed effects. Significance levels: p<0.10; p<0.05; p<0.01.To ensure comparability across specifications, both the logit and spatial bivariate probit models are estimated using the same sample of 3,244 observations. This sample is slightly smaller than that used in Table C.2 because observations located in FSAs without contiguous neighbors cannot be incorporated into the spatial adjacency matrix.



(a) Average $PM_{2.5}$ (b) Days $> 10\,\mu$g/m^3 (c) Days $> 27\,\mu$g/m^3

(d) Average $PM_{2.5}$ (e) Days $> 10\,\mu$g/m^3 (f) Days $> 27\,\mu$g/m^3

Figure C.1: Posterior distributions of the $PM_{2.5}$ coefficients. The upper row reports posterior distributions for low birth weight, while the lower row reports posterior distributions for preterm birth. The shaded region represents the posterior probability that the $PM_{2.5}$ coefficient is positive.

C.2 Spatial Bivariate Probit Model

We follow Neelon et al. (2014) and estimate a Bayesian spatial bivariate probit model to analyze the impact of $PM_{2.5}$ on preterm birth (PTB) and low birth weight (LBW) jointly. Let y_{1ij} denote an indicator equal to one if birth (i) in FSA (j) is preterm and y_{2ij} an indicator equal to one if the infant has low birth weight. The observed binary outcomes are generated by latent continuous variables:

$$y_{kij} = 1(y^*_{kij} > 0), \qquad k = 1,2. \tag{2}$$

The latent variables follow a bivariate normal distribution:

$$\begin{pmatrix} y^*_{1ij} \\ y^*_{2ij} \end{pmatrix} \Big| \phi_{\mathbf{i}} \sim N_2 \left[\begin{pmatrix} \eta_{1ij} \\ \eta_{2ij} \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right] \tag{3}$$

where ρ measures the within-individual correlation between PTB and LBW. The latent means are specified as

$$\eta_{kij} x'_{ij} \beta_k + w'_i \alpha_k + \phi_{ki}, \qquad k = 1,2, \tag{4}$$

where x_{ij} contains individual-level covariates, w_i contains FSA-level covariates, and ϕ_{ki} is a FSA-specific spatial random effect.

A key feature of the model is its treatment of spatial dependence. Defining $\boldsymbol{\phi}_i = (\phi_{1i}, \phi_{2i})'$, the spatial effects are assigned a bivariate intrinsic conditional autoregressive (BiCAR) prior:

$$\boldsymbol{\phi}_i \mid \boldsymbol{\phi}_{-i} \sim N_2 \left(\frac{1}{m_i} \sum_{l \in \partial_i} \boldsymbol{\phi}_l, ; \frac{1}{m_i} D \right), \tag{5}$$

where m_i denotes the number of neighboring FSAs and D is a 2×2 FSA-level covariance matrix:

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{pmatrix}. \tag{6}$$

The covariance term D_{12} captures the correlation between FSA-level PTB and LBW risks. A positive value implies that FSAs with unusually high PTB rates also tend to exhibit unusually high LBW rates after controlling for observed covariates.

The model therefore accommodates two distinct sources of dependence. The parameter ρ captures the correlation between PTB and LBW for the same infant, while D_{12} captures correlation in the spatial distribution of the two outcomes across FSAs. By combining a bivariate probit specification with correlated spatial random effects, the model permits joint, marginal, and conditional inference while accounting for both spatial clustering and outcome dependence.

We ran a single MCMC chain of 50,000 iterations, discarding the first 10,000 iterations as burn-in. Standard MCMC diagnostics, including trace plots, autocorrelation functions, and effective sample sizes, indicated good mixing of the model parameters, including the correlation parameter. Visual inspection of the trace plots suggested that the chain reached stationarity well before the end of the burn-in period. We also verified that thinning the chain had a negligible effect on the posterior summaries, and therefore report results based on the full set of post-burn-in draws.