Three Essays on Environmental Economics

by

Nusrat Jahan

A Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

of the University of Manitoba

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Economics

University of Manitoba

Winnipeg

Declaration

This thesis contains sole-authored work. I am the sole author of the text, analysis, and results presented in all chapters.

My supervisor, Professor Soodeh Saberian, provided guidance, feedback, and revisions throughout every stage, from developing the methodology to data analysis and writing all three chapters. Her insightful direction helped clearly define each chapter's contribution, refine the identification strategies, strengthen the empirical design, and improve the robustness of the analysis. Her constructive feedback on my writing enhanced the organization, coherence, and overall quality of the thesis.

The ideas for **Chapters 1 and 2** were developed collaboratively through extensive discussions with Professor Saberian. I framed the research questions, developed the methodologies, carried out the analysis, and prepared the initial drafts of these chapters under her supervision.

The idea for **Chapter 3** originated from Professor Saberian. I framed the research questions, developed the methodologies, conducted the analysis, and prepared the initial draft under her supervision.

Professor Saberian and I may pursue the publication of Chapters 1 and 2 in the future; however, at the time of submission, no formal publication plan is in place. If these chapters are published, I will be listed as the first author, with Professor Saberian recognized as the second author in acknowledgment of her significant contributions.

Abstract

This dissertation examines the effects of three distinct environmental stressors, including air pollution, flooding, and high temperatures, on mortality, educational attainment, and labor supply. It contributes to the broader environmental, health, and labor economics literature by providing causal evidence on previously understudied outcomes. Each chapter focuses on a context and population often overlooked, offering new insights into the human capital and productivity impacts of environmental shocks.

Chapter 1 estimates the causal impact of air pollution on early childhood mortality in Canada using a nationally representative death registry. Leveraging wind direction as an exogenous source of variation in daily $PM_{2.5}$ levels, I isolate the effect of both local and non-local pollution sources. A 1 $\mu g/m^3$ increase in fine particulate matter increases the early neonatal mortality rate per million births by 4.2 percent relative to the three-day mortality rate. Effects are most pronounced above 6.4 $\mu g/m^3$ $PM_{2.5}$ concentrations, supporting the case for more protective air quality standards.

Chapter 2 analyzes the cognitive effects of the 1997 Manitoba flood on children using standardized test data from the National Longitudinal Survey of Children and Youth. I measure flood exposure by distance to the flood polygon and apply a Difference-in-Differences strategy. A 1 km increase in distance is associated with a 0.4-points gain in test scores, with stronger effects for younger and medically vulnerable children. No income-based heterogeneity is observed.

Chapter 3 investigates how heat affects absenteeism among Indian service sector workers, using data from the Indian Human Development Survey-II. A 1°C increase in daily maximum temperature raises teacher absenteeism by 0.6 percentage points, with the effect concentrated in female-majority schools exposed to extreme heat (above 40°C). No significant

temperature effect is found for medical staff. Results are robust to alternative heat metrics, including diurnal temperature range and heat index.

Together, these chapters provide new evidence from Canada and India on the pathways through which environmental shocks affect human capital. The findings have clear policy implications ranging from strengthening air quality regulations and disaster preparedness to developing heat adaptation strategies, highlighting the need for environmental resilience in promoting public health, educational outcomes, and labor supply.

Acknowledgement

"So verily, with hardship comes ease. Indeed, with hardship comes ease."

Qur'an, 94: 5-6

It has been nearly seven years since I sailed this ship on a sunny afternoon. I glimpsed islands along the way and awaited the shore I was destined for. Today, I can see the shore; it is near, and I hope to reach it soon. All praise is due to my creator, the Lord of the worlds and heavens, who has guided me across this vast sea. Throughout my journey, I have learned, grown, and gained a lifelong moral that,

"We rise by lifting others."

— Unknown author

I want to express my deepest gratitude to my PhD supervisor, Dr. Soodeh Saberian, for her insightful guidance, unwavering support, and constructive criticism throughout this journey. I am sincerely indebted to my committee members, Dr. Julia Witt and Dr. Umut Oguzoglu, for their valuable feedback on my dissertation and their thoughtful, kind direction in completing this program. Ideas for my first and third chapters were developed in collaboration with my advisor, while the idea for the second chapter emerged from a discussion with my advisor and committee members. I am thankful to them.

I am forever grateful to my family for being my anchor; to my parents for their love and prayers, as well as for the sacrifices they made for my education; to my husband for his companionship, sacrifices, and ongoing support throughout this journey; to my son for every moment he missed my presence; and to my wonderful sisters for their prayers and lifelong support.

I am thankful to my sister, Israt Jahan, and our GIS librarian, Meg Miller, for their generous support and for introducing me to GIS. I feel blessed to have colleagues like Aslam, Thuhid, Iqbal, Atish, and Jennifer, who have consistently encouraged me on this journey and supported me when I needed it most. I want to express my gratitude to my teachers and professors who dedicated their valuable time to my education.

I am grateful to the Department of Economics for the resources and supportive learning environment I have experienced here. I also thank Statistics Canada's Research Data Center for providing data access for this dissertation. The experience I gained at the University of Manitoba is a priceless gem. To all my fellow sailors navigating the sleepless nights of high and low tides while holding their breaths, I wish you a bon voyage!

Dedication

This dissertation is dedicated to my mother, whose tireless work and unwavering support made my education possible, and to the three men in my life— my father, whose guidance has accompanied me every step of this journey; my husband, for his sacrifices, support, and constant encouragement; and my son, whose smile eased my stress and brightened even the hardest days.

Table of Contents

Introduction

Humans interact with the environment every day, starting from birth and even before in the mother's womb. Extreme environmental events driven by climate change and anthropogenic pollution affect individual well-being through multiple channels, impacting physical and mental health, cognitive function, and work performance. Although the adverse effects on individual living standards are more evident in the short term, the long-term cumulative impact on a country's human capital development and labor market outcomes is substantial. The three chapters of this dissertation contribute to the broader health, environmental, and disaster literature by examining unique research questions and providing new evidence across different geographical contexts relevant to policymakers and researchers interested in environmental health and climate change.

The first chapter presents evidence on how air pollution from industrial activities, wildfires, and transboundary sources affects public health in Canada, despite strict federal air quality standards nationwide. Specifically, we examine whether short-term exposure to $PM_{2.5}$ has a causal effect on early childhood mortality in Canada. We analyze non-accidental death events of children under five years old using a nationally representative administrative death dataset from Statistics Canada. To address endogeneity arising from reverse causality, measurement errors, omitted-variable bias, or locational sorting, we use wind-based instruments to predict exogenous variation in daily pollution exposure. The Instrumental Variable (IV) method estimates the effect of non-local pollution exposure on three-day total child mortality at the census division level. Our results show that a 1 $\mu g/m^3$ increase in fine particulate matter raises mortality rates per million births by about 1.0 across early neonatal, neonatal, infant, and under-five death categories over three days. Our nonlinear estimates indicate significant effects for $PM_{2.5}$ levels exceeding the average concentration of 9 $\mu g/m^3$.

The findings recommend that air quality policies at the provincial and federal levels in Canada be improved to safeguard vulnerable populations.

In the second chapter, we examine the causal effect of flood exposure on children's cognitive abilities in Manitoba by using standardized test scores from the National Longitudinal Survey of Children and Youth. To our knowledge, this is the first study in Canada, North America, or any developed-country setting to investigate the causal effect of flood exposure on children's standardized test scores, despite the increasing frequency of flood events. We measure flood exposure by calculating the distance between the 1997 overland flood polygon and the centroids of each Forward Sortation Area. We then employ a Difference-in-Differences method to estimate the causal effect of flood exposure, using distance to the flood polygon as a continuous treatment variable. The results indicate that a 1 km increase in distance to the flood polygon is associated with a 0.4-point increase in average test scores. The findings suggest that flood exposure affects children's test scores across all age groups, with girls being more vulnerable than boys. We also find that learning disabilities, when interacting with flood exposure, negatively affect test scores. Conversely, subsample analyses based on birth prematurity, long-term health condition, or household income status show no significant effects. The key contribution of this chapter is to provide evidence that even high-income countries with better infrastructure, resources, disaster preparedness, and emergency response plans are not completely immune to the effects of extreme weather events.

In the final chapter, we investigate the relationship between temperature and service-sector absenteeism. The effects of outdoor and indoor temperatures on labor supply and productivity are well-documented in the literature. But most studies rely on aggregate economic output or on industrial sectors, leaving service sectors largely unexamined despite their essential role in human capital formation. We fill this gap in the literature by examining the causal effect of ambient temperature on employee absence rates in India, using data from the Indian Human Development Survey-II (IHDS-II). We analyze the education and health sectors to provide evidence that temperature effects could be highly sector-specific. The average district-level teacher absence rate in primary schools is 12.53%, while the average daily maximum temperature is 31°C. By estimating a fixed-effects model with weather and school controls, we show that a 1 °C increase in temperature leads to a 0.6% rise in teacher absence rates. This effect is more pronounced in schools with a higher proportion of female teachers, particularly in areas with daily maximum temperatures exceeding 40 degrees Celsius. The estimates do not suggest any lagged effect of temperature on teacher absence rates. Additionally, the estimates remain consistent when alternative temperature measures, including the Diurnal Temperature Range (DTR) and the heat index, are considered. Conversely, temperature appears to have no effect on the medical staff absence rate, possibly due to higher job responsibilities, better work environments, or improved salary structure.

Synthesizing findings across all three chapters, it is clear that environmental degradation caused by climate change affects human capital development at different stages of life, and that all countries face risks regardless of their economic status or resource endowment. Our findings highlight the need for local and national policies, including strengthening air quality standards, improving disaster preparedness programs, and creating heat-adaptation strategies, to mitigate the effects of environmental shocks on human health and well-being.

Chapter 1 Air Pollution and Early Childhood Mortality: Evidence from Variations in Wind Direction in Canada

Abstract

Using the nationwide Canadian Vital Statistics - Death (CVSD) database, we provide compelling evidence of the significant causal impact of air pollution on all-cause child mortality rates. To identify causality, we exploit exogenous variations in fine particulate matter driven by changes in daily wind direction. Our findings indicate that a 1 $\mu g/m^3$ increase in fine particulate matter, resulting from specific wind patterns, raises the early-neonatal mortality rate by 4.2 percent relative to the three-day mortality rate. We find no evidence of mortality effect in females, while males are particularly vulnerable across all categories of death. Our results are robust and withstand rigorous sensitivity tests.

Keywords: Air pollution – mortality – social costs.

1 Introduction

This paper conducts a large-scale, quasi-experimental investigation of the short-term effects of fine particular matter ($PM_{2.5}$) exposure on all-cause child mortality in Canada, providing the first evidence of *causal* impacts of $PM_{2.5}$ exposure on the childhood mortality rate in the country.

Well-documented negative impacts of exposure to $PM_{2.5}$ during pregnancy on birth outcomes (Ng et al., 2017; Sheridan et al., 2019; Wang et al., 2018; Zang et al., 2019) underscore the vulnerability of children to indirect exposure to $PM_{2.5}$. Adverse effects of air pollution exposure during pregnancy include later respiratory morbidity, low birth weight, impaired lung function, and high infant mortality (Proietti et al., 2013).

Numerous epidemiological and environmental studies indicate a substantial association between $PM_{2.5}$ and neonatal mortality (Goyal et al., 2019), as well as post-neonatal mortality attributable to respiratory causes (Woodruff et al., 2006), infant mortality (Bachwenkizi et al., 2021; Heft-Neal et al., 2020), and mortality in children under five years of age (Owili et al., 2017). For instance, Ortigoza et al. (2021) estimates an association between $PM_{2.5}$ and mortality in children under five and infants in 337 cities across Latin America. The authors discovered a 1 μg/m3 increase in $PM_{2.5}$ corresponds to a 0.4% increase in under-five mortality and a 0.5% increase in infant mortality. Nevertheless, they reported no association between $PM_{2.5}$ and mortality in children aged 1 to 4. Drawing on data from 58 developing countries, Hassan & Murad (2022) concluded that a 1% increase in outdoor $PM_{2.5}$ concentration results in a 0.2% increase in respiratory-related child mortality. Their findings indicate that outdoor $PM_{2.5}$ significantly exacerbate post-neonatal mortality, whereas neonatal mortality is predominantly linked to indoor air pollution. Yorifuji et al. (2016) examine the relationship between $PM_{2.5}$ and the mortality data of 2,086 infants from 23 urban areas in Japan. The authors reported that each 10 μg/m^3 increase in $PM_{2.5}$ is associated with an elevated risk of post-neonatal and infant mortality rates, even when $PM_{2.5}$ concentration remains below Japan's Air Quality Standards.

To our knowledge, this paper provides the first evidence of the *causal* effect of acute $PM_{2.5}$ exposure on all-cause mortality in children in Canada. Our study broadens the boundary of negative impacts of air pollution by pursuing two main objectives: First, to causally link daily variations in $PM_{2.5}$ to children's mortality in Canadian cities. Second, to investigate the mechanisms underlying our findings by testing the sensitivity of our results to the confounding role of the climatic factors.

The significance of this issue is particularly noteworthy considering the reputation of Canadian cities for their overall good air quality. Despite this reputation and contrary to expectations, there has been an increase in particulate pollution levels in Canada over the past decade (Figure 2). The annual mean levels of $PM_{2.5}$ have experienced a rise from 6 µg/m³ in 2008 to 7.2 µg/m³ in 2018. On the other hand, the national infant mortality rate has shown a fluctuating pattern, falling from 5.3 per 1,000 live births to 4.5 per 1,000 live births in 2015 and then climbing to 4.8 per 1,000 live births in 2022.[1] This upward trend in particulate pollution and mortality rate highlights the need for a closer examination of potential implications for public health, including the impact on children's mortality. Investigating the mortality effects of air pollution in Canada enables us to gain a comprehensive understanding of the health risks associated with air pollution, even in regions known for their good air quality.

Estimating the causal effects of air pollution on children's mortality is complicated due to well-documented challenges, including residential sorting, confounding factors, and measurement errors in assigning pollution exposure. Existing quasi-experimental studies that rely on exogenous fluctuations in air pollution as their primary sources of identification have

[1] National infant mortality rate at Statistics Canada
https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310071301

typically been confined to narrow geographic regions, limiting their external applicability. Furthermore, due to constraints in available data, these studies often struggle to identify the impacts on critical yet infrequent outcomes such as children's mortality.

To address the estimation challenges, we employ comprehensive administrative data from the Canadian Vital Statistics - Death Database (CVSD) and daily climatic and pollution data from 2006 to 2018 to exploit exogenous variations in $PM_{2.5}$ concentrations resulting from changes in daily wind direction. The identifying assumption, which is very similar to Deryugina et al. (2019), relies on the fact that after flexibly controlling for climatic variables, spatial and temporal fixed effects, changes in daily wind direction are exogenous to changes in children's mortality, except through their impacts on $PM_{2.5}$ levels. Compared to the previous quasi-experimental works, our study uses variations in $PM_{2.5}$ exposure across a wide geographic area and over an extended timeframe.

We estimate that one microgram per cubic meter (1 $\mu g/m^3$) about (0.1 standard deviations) increase in daily $PM_{2.5}$ exposure increases early neonatal mortality by 1.0 over the three-day window that spans the day of the increase and the following two days. As demonstrated, our IV estimates are higher than both the ordinary least squares (OLS) and correlational estimates reported in the observational studies, demonstrating the potential for substantial bias in those studies. The effect sizes, however, are higher than those in Farhat et al. (2013), compared to their estimates of death for people of all ages in 12 Canadian cities, and Anwar et al. (2021), compared to their estimates of mortality for children under the age of five in 16 Asian countries. Conversely, the effect is relatively lower than those in Alnwisi et al. (2022) compared to their annual estimates of under-five mortality in China.

Our estimates show that infants do not suffer from 'mortality displacement'.[2] Following Deryugina et al. (2019), we show that the effect of one-day $PM_{2.5}$ exposure on child mortality grows if we expand the time window over which we measure mortality from 1 to 3 days. However, it shows a fluctuating pattern when we further increase the time window to 5, 7, 10, and 14 days, suggesting a possible next episode of higher mortality rate in the risk pool.

In 2021, after 15 years, the World Health Organization (WHO) tightened its recommendation for $PM_{2.5}$, considering only the negative cardiovascular impacts of air pollution exposure. The results from our work point to the fact that even current assessments of the social costs of pollution exposure underestimate the benefits of clean air and, by implication, prescribe regulatory standards that are too lax. The strand of research to which this study contributes that air quality improvements can enhance health outcomes has potentially game-changing implications for evaluating the benefits generated by policy interventions to reduce air pollution.

Our results may also lead us to re-assess the prevailing wisdom that policy promulgation requires a trading-off of health benefits and financial costs while there may be no such trade-off. At a grander level, we hope to contribute another pellet of evidence to the growing view that there is no inexorable conflict between environmental outcomes and economic success as has traditionally been supposed. Our analysis suggests that a cleaner city is a healthier city with a lower children's mortality rate. The rest of the paper is organized as follows. Section 2 provides the mechanism by investigating previous studies. Section 3 describes our data. Section 4 explains our empirical strategy in detail. Section 5 presents the results, and Section 6 concludes.

[2] Deaths that occur sooner than they otherwise would are called mortality displacement. This occurs when deaths move from the future into the present, changing a person's life expectancy.

2 Literature review

Previous research has documented the associations between $PM_{2.5}$ exposure and child morbidity and increased hospitalization rates due to acute respiratory illnesses (Xia & Yao, 2019; Dunea et al., 2016; Davila Cordova et al., 2020). These associations also extend to childhood pneumonia, acute bronchitis, and asthma (Tecer et al., 2008; Ostro et al., 2009; Egondi et al., 2018).[3]

For example, Egondi et al. (2018) identified a correlation between $PM_{2.5}$ and child morbidity in the urban slums of Kenya, particularly in regions characterized by elevated pollution levels. Their findings indicated that children residing in areas with significant $PM_{2.5}$ concentrations are 25% more likely to experience symptoms such as cough, fever, or convulsions and have a 12% higher likelihood of mortality due to respiratory-related illnesses. Additionally, the mortality rate in highly polluted areas is markedly elevated among children under the age of one, particularly within poor households. Furthermore, various studies have established links between early childhood exposure to $PM_{2.5}$ and other health indicators, including childhood height-for-age (Spears et al., 2019), childhood overweight (Mao et al., 2017), and lung function in children of school age (Zwozdziak et al., 2016).

The well-known in-utero effects of $PM_{2.5}$ on birth outcomes (Ng et al., 2017; Sheridan et al., 2019; Wang et al., 2018; Zang et al., 2019) indirectly show how susceptible children are to direct $PM_{2.5}$ exposure right after birth and in the early years of their lives. We aim to study effect of $PM_{2.5}$ on children at a very early stage of life, between as early as one day of age and five years of age. In this section, we first document the health effect of $PM_{2.5}$ in children and adults across Canada. Next, we discuss the biological process that underlies children's susceptibility to $PM_{2.5}$ exposure.

[3] We narrowed our literature search to only the health effects of $PM_{2.5}$ rather than considering all criteria pollutants.

2.1 Health effects of $PM_{2.5}$ in Canada

The detrimental health effects of $PM_{2.5}$ in Canada have garnered substantial attention in recent health and epidemiological research. Numerous studies have estimated the mortality risk attributed to even low levels of $PM_{2.5}$ (Chen et al., 2022; Shin et al., 2021; Christidis et al., 2019; Pappin et al., 2019; Cakmak et al., 2018; Pinault et al., 2016; Elliott & Copes, 2011). Additional research has indicated associations between $PM_{2.5}$ and physical morbidity, including diabetes (Pinault et al., 2018; Crouse et al., 2015), asthma and respiratory illnesses (Stieb et al., 2019; Matz et al., 2020), hypertension (Requia et al., 2017), lung cancer (Gogna et al., 2019), and mental disorders (Weichenthal et al., 2016).

Among these studies, Shin et al. (2021) investigate the impact of ambient $PM_{2.5}$ exposure on hospitalization and all-cause mortality in Canada. Their findings reveal that short-term exposure (a 2-week exposure time window, 0, 1, …13) to $PM_{2.5}$ correlates with increased mortality rates among women, while higher hospitalization rates for men. In another study, Crouse et al. (2015) analyze the relationship between $PM_{2.5}$ and cause-specific mortality in Canada. Their results indicate that $PM_{2.5}$ is connected to respiratory, cardiovascular, and diabetes-related mortality among respondents aged 25 and older.

Further research has addressed the health effects of $PM_{2.5}$ on Canadian children, identifying issues such as respiratory illnesses (Lavigne et al., 2021; Wong et al., 2016), early childhood cancer (Lavigne et al., 2017), and adverse birth outcomes (Stieb et al., 2016; Stieb et al., 2019). In Ontario, To et al. (2021) identify associations between components of $PM_{2.5}$ and all-cause acute care hospitalizations in children aged 4 to 9 years. However, their study found no impact of $PM_{2.5}$ on acute hospitalizations related to asthma and associated disorders. Another investigation by Wong et al. (2016) confirms that industrial $PM_{2.5}$ emissions are associated with airway obstruction and compromised lung function in male children, while no such correlation is observed in female children.

2.2 Biological mechanisms and child mortality

Most epidemiological research has linked postnatal exposure to air pollution to poor health outcomes. When compared to adults, newborns breathe in more polluted air, which accelerates the development of respiratory and lung disorders at this time and stalls lung function until adolescence (Zhao et al., 2021). According to Milani et al. (2022), newborns' short-term bronchiolitis symptoms are substantially associated with their exposure to $PM_{2.5}$, for the first 29 days after birth. However, low-to-moderate $PM_{2.5}$ exposure alters the nasal microbiome in newborns, which can result in the development of asthma (Gisler et al., 2021), and it also has an impact on a baby's gut health by changing the makeup of bacteria at six months of age (Bailey et al., 2022). Furthermore, the Mental Development Index and Psychomotor Development Index scores at age two are said to decline by 3.4 and 10.7 points, respectively, for every 10 g/m^3 increase in postnatal $PM_{2.5}$ exposure, according to Wang et al. (2022). Also, children aged 0 to 3 who have been exposed to $PM_{2.5}$ early in infancy experienced sleep difficulties (Cai et al., 2023). Additionally, research on mice and humans suggests that $PM_{2.5}$ may hasten the onset of atherosclerosis or the thickening of the arteries brought on by a buildup of plaque in the inner lining of an artery, which can be fatal (Künzli et al., 2005; Sun et al., 2005).

A time series study on 12 Canadian cities finds that a 10 g/m^3 increase in the 24-hour average $PM_{2.5}$ concentration is associated with a substantial increase in all-age mortality ranging from 0.9% to 3.2% (Farhat et al., 2013). Likewise, in sixteen Asian countries, Anwar et al. (2021) found that a one-unit annual increase in $PM_{2.5}$ causes a nearly 14.5% increase in the number of children dying before age five. They arrive at this conclusion by using a two-stage least squares approach that takes advantage of variations in $PM_{2.5}$ attributable to economic growth in estimating the effect on child mortality. According to He et al. (2022), newborns in China under 28 days of age experience an increase in all-cause mortality of 1.5

percent because of short-term exposure (lag 0 to 1 day) to $PM_{2.5}$. They also found association between under-5 mortality and $PM_{2.5}$ exposure for several major causes such as pre-term births, low birth weight, pneumonia, birth asphyxia, congenital abnormalities, diarrhea, and digestive disease. However, in a national analysis in China, Alnwisi et al. (2022) reports that under-five mortality rises by 1.2 per 1000 live births for every 10 g/m^3 increase in yearly $PM_{2.5}$ concentration. Also, observational research using US vital records Khadka & Canning (2022) reveals that exposure to $PM_{2.5}$ is positively correlated with infant mortality in the first and second months after birth.

None of these papers except He et al. (2022) concentrates on neonatal mortality or intensively address the effects of $PM_{2.5}$ on early neonates, neonates, and post-neonates, although they provide a solid foundation for researching child mortality below five years of age.

2.3 Contribution

While Canadian cities generally maintain better air quality and meet federal standards, it remains essential to assess how short-term exposure to fine particulate matter ($PM_{2.5}$) from industrial activity, wildfires, or transboundary pollution affects population health. Existing Canadian studies link $PM_{2.5}$ to adult physical and mental health, all-cause or cause-specific mortality, and childhood respiratory morbidity, but none examine early childhood mortality attributable to $PM_{2.5}$. Internationally, most studies on neonatal, post-neonatal, and infant mortality document associations rather than causal effects, with few exceptions such as Chay and Greenstone (2003a; 2003b), who use quasi-experimental designs to show that declines in total suspended particulates in the U.S. significantly reduced infant deaths, particularly within the first month of life. This chapter fills that gap by providing the first causal evidence for Canada on how $PM_{2.5}$ exposure affects mortality across different childhood age groups up to age five, using a nationally representative administrative death dataset. Even in a low-pollution context with stringent air quality standards, we find statistically significant and policy-relevant health impacts, underscoring the vulnerability of young children to low-level $PM_{2.5}$ exposure. The results have important economic and environmental policy implications: they highlight the need to consider environmental vulnerability even in low-risk settings, account for the long-term effects on Canada's human capital, and inform revisions to air quality standards and public health guidelines. These findings also establish a baseline for international comparisons with countries facing more severe air quality challenges.

3 Data

3.1 Child mortality

We prepare child mortality data from all death information presented in the Canadian Vital Statistics Death Database (CVSD) from 2006 to 2018. The CVSD is an administrative database that provides historical data on death events, demographics, and the geographical location of the deceased.[4] We aggregate all non-accidental death events in a Census Division (CD) to the CD-date level using the deceased's residence census division.[5]

Table 1 displays brief statistics for the 500,310 CD-Date observations that comprise our primary estimate sample.[6] We define early neonatal death as a death occurring under one week of age (0 to 6 days), neonatal death as a death occurring under four weeks of age (0 to 27 days), post-post neonatal death as a death occurring under one year of age but at least 28 days old (28 to 364 days), and infant death as a death occurring under one year of age (0 to 364 days)

---

[4] The CVSD provides geographic code for both place of death census division and residence census division of the deceased following Standard Geographical Classification (SGC) 2011. See https://www150.statcan.gc.ca/n1/pub/12-571-x/12-571-x2011001-eng.pdf

[5] Our study emphasizes the under-five age group (0 days to below 5 years) for which the residence census division represents the place of death census division, with more information available for the former.

[6] Our sample covers all ten provinces in Canada, excluding the three territories.

following the guidelines in the CVSD user guide. We define one-to-five death as a death occurring at or above 365 days and below five years old. Lastly, we define under-five death as a death occurring under five years of age (0 to under 1825 days or 60 months) following Wagstaff et al. (2007). There are, on average, 38 under-five death events per million in the sample, of which 21 are early neonatal deaths, 25 are neonatal deaths, and 33 are infant deaths. Number of deaths is higher for boys compared to girls across all death categories.

By the United Nations standards and indicators guide, we compute the mortality rate for all death categories using the following formula and express it as a rate per million.[7] We obtain CD-level annual live births and under-five population data from Statistics Canada's demographic estimate files for the 2016 census.[8]

i) Early neonatal, Neonatal, and Infant

$$\text{mortality rate} = \frac{\text{Number of deaths in the respective death category}}{\text{Total number of live births}} \times 1{,}000{,}000$$

ii) $$\text{Under-five mortality rate} = \frac{\text{Number of deaths among under fives}}{\text{Total under five population}} \times 1{,}000{,}000$$

3.2 Climate

It is sensible to consider various meteorological elements as potential confounders since they might affect mortality independently. It is also known that climatic conditions affect pollution concentration. We retrieve hourly data on weather variables such as temperature, precipitation, humidity, wind speed, and wind direction from Environment and Climate Change Canada (ECCC).[9] ECCC provides surface meteorological data on these weather parameters observed at each aerodrome reference point across Canada.

---

[7] See United Nations' standards and indicators guide (2010) at https://www.unhcr.org/media/53-standards-and-indicators-guide-revised

[8] Census data search at https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710013901

[9] ECCC data search at https://climate.weather.gc.ca/historical_data/search_historic_data_e.html

Data on wind velocity (direction and speed) are expressed as tens of degrees and kilometers per hour and represent winds 10 meters above the ground. We convert the 10's degree hourly wind direction to arc degree and then construct a daily measure using an arctangent function.[10] The prevailing wind direction in Canada is from the west, but it changes due to the local variations in wind pressure, temperature, and landscapes. Therefore, we define wind direction as the direction from which the air blows.[11]

We also obtain cloud cover data from the Automated Surface Observing System (ASOS) of Iowa Environmental Mesonet, which maintains airport weather observations worldwide. Using the same computation as for pollution, we transform all hourly weather information into daily values to construct a CD-date measure for each census division.

3.3 Air pollution

We obtain data on ambient air pollutants from the Canada-Wide Air Quality Database of the National Air Pollution Surveillance System (NAPS) from 2006 to 2018.[12] NAPS has provided uniform-standard data on all criteria pollutants since 1972, including fine particulate matter $PM_{2.5}$ starting in 1995. We collect data on $PM_{2.5}$ and O_3 since previous literature asserts significant adverse effects of these pollutants on mortality and birth outcomes (Deryugina et al., 2019; Ranjbaran et al., 2020; Yang et al., 2018; Zang et al., 2019). We convert hourly readings to daily measures for each monitoring station and then take a daily average of all stations in a CD to generate pollution measures at the CD-date level.

[10] Implies that a full rotation of a circle is 360 degrees or 2π radians where one degree is equivalent to $\frac{\pi}{180}$ radians.

[11] 90 degrees means an east wind and 360 degrees means a wind blowing from the geographic North Pole. See https://climate.weather.gc.ca/glossary_e.html

[12] NAPS data search at http://data.ec.gc.ca/data/air/monitor/national-air-pollution-surveillance-naps-program/

Figure 2 presents the annual mean concentration of $PM_{2.5}$ at CD level during our study period and the annual number of NAPS stations across the ten provinces in Canada. The upward trend in the first panel shows that the average concentration of $PM_{2.5}$ steadily rose from 2006 to 2018, with a sharp fall in 2016. Conversely, the number of stations has remained roughly constant, with seven more added between 2008 and 2011. This is concerning since Canada consistently ranks high among the countries with the cleanest air in the world. However, $PM_{2.5}$ has an average daily concentration of 6.4 $\mu g/m^3$ with a standard deviation of 10.1 over the sampling period (See Table 1).

4 Method

We aim to investigate short-term causality between fine particulate matter and all-cause early childhood mortality, excluding any possible confounding influence. Therefore, we estimate the following model using Two Stage Least Square (2SLS):

$$ECMR_{CDdmy} = PM_{2.5\,(CDdmy)}\theta + W_{CDdmy}\gamma + FE_{CD} + FE_{pm} + FE_{my} + \epsilon_{CDdmy} \quad (1)$$

where $ECMR_{CDdmy}$ is the early childhood mortality rate for each death category in census division CD on date d in month m and year y. The coefficient θ measures the effect of an extra unit of $PM_{2.5}$ on the mortality rate on a typical day. Following Deryugina et al. (2019), we use a three-day total mortality rate measure to account for potential mortality displacement in our sample. For example, on a typical exposure day t, our dependent variable is the sum of mortality rates on day t, (t + 1), and (t + 2). This means any death event on day t does not have any effect on the mortality rate [t + (t + 1) + (t + 2)]. Therefore, θ exclusively captures the effect of an increase in daily $PM_{2.5}$ on the child mortality rate, net of any modification caused by mortality displacement.

We add a vector of weather variables W_{CDdmy} to control the confounding effect of climate conditions. This vector includes the daily averages of temperature, precipitation, humidity, wind speed, and cloud cover. We construct 25 bins for the temperature range between -36 degrees Celsius (-32° F) or less and 33 degrees Celsius (91.4° F) or more, where each of the 23 intermediate bins spans 3 degrees Celsius (37.4° F). Then, we generate a factor variable for temperature grouping for all 25 temperature bins. We add a precipitation indicator variable, create a dummy for each decile of wind speed, and add it as a factor variable in the model.

Next, we include census division (FE_{CD}), province-by-month (FE_{pm}) and month-by-year (FE_{my}) fixed effects. FE_{CD} captures unobserved spatial heterogeneity in child mortality, child health, and pollution concentrations at the census division level. By adding province-by-month fixed effects, we allow for provincial variation in child health and the seasonal correlation between pollution and wind direction. Lastly, month-by-year fixed effects capture any time-varying policy change in health care or pollution regulation during the study period. All regressions include clustered standard errors at the CD level. Our estimates are robust to alternative choices of fixed effects and individual weather variable selections.

Ordinary Least Square (OLS) estimate of equation (1) would result in biased estimates as $PM_{2.5}$ is not randomly assigned because confounding weather conditions might result in omitted variable bias (Neidell, 2004), while inadequate information on an individual's exposure to pollution at different times on a particular day might cause measurement error (Graff Zivin & Neidell, 2009). Moreover, residential sorting might arise if high-income groups choose to live in a less polluted area for a healthier life (Graff Zivin & Neidell, 2009; Greenstone & Chay, 2005). We employ an instrumental variables (IV) approach, using daily wind direction at the census division level as an instrument for $PM_{2.5}$ concentrations. To account for spatial heterogeneity, we allow the effect of the wind-based instruments on $PM_{2.5}$ to vary by geographic cluster. The first-stage specification is as follows:

$$PM_{2.5\,(CDdmy)} = \sum_{k\epsilon K=2}^{17} \sum_{d=2}^{36} \beta_{dk} (C_k \times Windbin_{CDdmy}^{10d}) + W_{CDdmy}\delta \qquad (2)$$

$$+FE_{CD} + FE_{pm} + FE_{my} + \epsilon_{CDdmy}$$

In the first stage, we calculate the effect of wind direction on $PM_{2.5}$, allowing estimates to vary by geography. In equation (2), our preferred instrument is the interaction between CD clusters and wind direction bins. The factor variable C_k indicates that census division C belongs to a cluster k from the set of station groups K. $Windbin_{CDdmy}^{10d}$ is a factor variable with 36 categories, each representing a wind bin with a 10-degree interval $[10d, 10d + 10)$.[13] Our parameter of interest, β_{dk} captures the geographically varying interaction effects of the instruments on daily mean $PM_{2.5}$ concentration at the CD-date level. Other variables, including all flexible weather controls and fixed effects, are as indicated in equation (1). All regressions include standard errors clustered at the CD level. Figure 4 and Figure 7 display the results from our first-stage estimation.

In order to assign pollution exposure to each census division (CD), we follow Deryugina et al. (2019) to group NAPS (National Air Pollution Surveillance) stations and, by extension, the CDs in which they are located into 17 clusters (Figure 3) using an unsupervised machine learning algorithm, the *k-means* clustering (Macqueen, 1967).[14] We initially run the k-means algorithm with K=10, 35, and 100. All three converge and show that the Within-clusters Sum of Squares (WSS) trend flattens after cluster 4, suggesting 4 clusters as the optimal number under the standard elbow method. However, grouping all NAPS stations in Canada into only 4 clusters does not yield sufficient variation in our treatment variable, which relies on wind

[13] Starting with $0^{\circ}(360^{\circ})$, our wind bins are $[0^{\circ}, 10)$, $[10^{\circ}, 20)$, $[20^{\circ}, 30^{\circ})$, … … …, $[350^{\circ}, 360^{\circ})$.

[14] K-means clustering creates groups of similar NAPS stations by identifying similarities through an iterative process. This algorithm consists of three steps. First, we specify the initial value of the parameter K and randomly assign each NAPS station to a group or cluster. Second, the algorithm calculates the centroids for each cluster. Third, it reassigns each station to the nearest cluster based on the computed distance between each station and each centroid. Finally, the process repeats steps two and three until convergence, indicating that the centroids and station assignments remain stable and do not change further. To determine the optimal number of clusters, we look for an elbow, which represents a significant drop in the distribution of Within-clusters Sum of Squares (WSS) (Thorndike, 1953).

direction shifts to generate plausibly exogenous changes in $PM_{2.5}$. To address this, we relax the standard "elbow" choice and instead select the point on the WSS curve where the sum of squared errors is minimized. This yields 17 optimal clusters when starting from K = 100 and running 20 iterations. This finer clustering reflects the fact that several provinces contain more than one NAPS cluster, as our sample spans all ten provinces.

Clustering in this way is important for measurement accuracy. Pollution monitor readings in a single CD may not represent the average exposure of all residents due to sparse and uneven monitor placement. By grouping monitors into geographically coherent clusters and constraining the effect of wind direction on pollution to be the same for all monitors within a cluster, we reduce the influence of highly localized emission sources that would otherwise generate measurement error. Local sources tend to affect monitors differently depending on their position relative to the source, whereas nonlocal pollution transported by prevailing winds tends to have a more uniform impact across all monitors in a cluster. This makes the wind-based variation more representative of actual area-wide exposure and less prone to bias from unrepresentative local fluctuations.

For example, if a local source is located in the centre of a CD and the monitor is to the east of it, westerly winds will raise measured pollution levels while easterly winds will lower them—yet in both cases, only part of the CD experiences the change. Aggregating exposure at the CD level without accounting for this leads to measurement error and potential bias in estimated health effects. By clustering monitors into broader regions, we capture pollution variation that is more likely driven by nonlocal sources, reducing this bias.

In our IV framework following Deryugina et al. (2019), we use CD-level wind direction as an instrument for $PM_{2.5}$, estimating first-stage effects that vary by geographic cluster. This approach leverages exogenous variation from wind-driven pollution transport while mitigating errors arising from sparse monitor coverage and heterogeneous local source impacts.

Pollution monitors in a particular census division experience varying degrees of wind direction because of a scattered distribution. Consequently, we suspect that pollution readings from these monitors might fail to represent the actual pollution exposure for those CD residents. Our interaction instrument in equation (2) ensures that effect of CD wind direction on average $PM_{2.5}$ is the same for all monitors in a geographic cluster. Moreover, different monitors in a group may record diverse pollution readings from a local source depending on their relative proximity to that source. Conversely, non-local pollution sources have identical effects on all monitors in a group as they are geographically located on one side or the other of the entire monitor group. Therefore, our estimates in equation (1) plausibly account for variation from non-local pollution sources.[15] This is expected in our setting because the restriction on cluster-level wind direction allows us to assign pollution in a better way and removes anomalies induced by the local source of pollution- for example, overestimating or underestimating individuals' pollution exposure since local pollution affects them differently.

Figure 1 depicts a wildfire incident in McBride or Red Deer Creek during the wildfire season from April to September. Consider that we want to assign $PM_{2.5}$ generated from this wildfire to residents in the CD, Fraser-Fort George, which includes Mackenzie, McBride, and Valemount. Any NAPS monitor close to Mackenzie will record a high level of $PM_{2.5}$ when wind blows from the south-east and a low level of pollution when wind blows from the north-west. Conversely, Any NAPS monitor close to Valemount will record a high level of $PM_{2.5}$ when wind blows from the north-west and a low level of pollution when wind blows from the south-east. Either way, half of the CD population experiences higher pollution, which is not a representative pollution measure for the entire census division. Moreover, using such variation

[15] Since all NAPS stations are decently scattered in a census division, we rationally disregard the presence of any potential pollution clusters on either side of the monitor groups. However, we have conducted additional exercises reducing the size of monitor groups to 15 and 6 instead of 17 to check if covering a larger area in a cluster affects our main estimates. The results in Table 4 support our preferred specification.

may result in spurious estimations suggesting that $PM_{2.5}$ does not significantly affect mortality if pollution from both halves offsets each other completely.

However, all monitors in a census division will record the same level of $PM_{2.5}$ when migrated from a non-local source. For instance, a predominant south wind during the summer will transport $PM_{2.5}$ away and evenly distribute it to Fort St. John and Fort Nelson, which are located in the same census division. Likewise, Edson, Hinton, and Yellowhead, which are in the same CD, will have identical $PM_{2.5}$ exposure with a predominant west wind.

Therefore, wind direction is constant for all 10-degree wind bins in our instrument to each monitor cluster. We employ the maximum number of wind bins to have a wide range of variations in wind direction over the sampling period. As a robustness check, we estimate equation (1) with 90-degree and 45-degree wind bins, and the results are close to our main specification.

5 Results

Columns (1) to (4) in Table 2 present the effects of daily $PM_{2.5}$ on child mortality rates for different death categories based on their age at death. Panel A lists negative OLS estimates for all death categories, implying that a 10-unit additional increase in daily $PM_{2.5}$, which is approximately equivalent to one standard deviation, lowers the three-day mortality rate, which is only significant for infant and under-five child mortality rate over the three-day window. Since OLS estimates are prone to bias, we report corresponding IV estimates in Panel B to show the causal effect of $PM_{2.5}$ exposure on the three-day mortality rate. These estimates are positive for all death categories and significantly higher than OLS estimates except for one to five. Column (1) suggests that each 1 $\mu g/_{m^3}$ increase in daily $PM_{2.5}$ exposure increases the three-day early-neonatal mortality rate per million births on average by 1.0, which is approximately 4.2 percent relative to the three-day average mortality rate. Columns (2) to (6) show that the corresponding estimated coefficients for neonatal, infant, and under-five populations are 1.0, 1.0, and 0.2, considering a 1 unit rise in the daily average concentration of $PM_{2.5}$. Therefore, on average, the mortality rate per million increases almost at the same rate in all death categories, with a relatively lower effect on the under-five age group. The effect relative to the mean of mortality rates is higher for early neonates and neonates (columns 1 and 2) compared to infants and under-five (columns 4 and 6). This implies that early neonates and neonates are more vulnerable to air pollution until they reach their first birthday compared to children above one year who subside the overall effect of $PM_{2.5}$ on the under-five age group.

Following the consistent results from Table 2, we estimate the three-day mortality effect of $PM_{2.5}$ separately for male and female children. Panel A of Table 3 shows that IV estimates are slightly higher for males in all death categories than our results in Table 2. However, the short-term variation in $PM_{2.5}$ has no significant effect on females in any death category. Thus, our main results in Table 2 fundamentally portray the effects on males, balancing out the insignificant effects on females.

We intend to compare our results with research on the detrimental health effects of air pollution or the health advantages of controlling air quality for children. Most epidemiological research estimates the effect of $PM_{2.5}$ on morbidity or cause-specific mortality for the elderly. A few studies discuss the effects of criteria pollutants on pregnancy and birth outcomes, which makes a direct comparison difficult. To our knowledge, this paper is the first to study the causal link between air pollution and child mortality by different age groups. However, we compare our estimates to Deryugina et al. (2019) as we follow their methodological approach in our settings. Using health information of more than 97 percent of US elderly, they report that a 10 $\mu g/m^3$ increase in daily $PM_{2.5}$ leads to 0.7 additional deaths in elderly mortality relative over the three-day window for the age group 65+ and 0.6 additional deaths for the age group 75+ for the one-day window. Our estimates for the three-day mortality rate (Table 2) are higher than those for the single-day mortality rate (Table 7), consistent with the abovementioned mortality pattern. Again, our same-day estimate increases over the second and third days for early neonates and neonates and over the fourth day for infants and children aged under five. Once again, this implies that children are relatively more susceptible to environmental pollution in the first few days of their birth.

Early-day literature suggests the effect of air pollution, in general, is distributed over time, i.e., it continues to affect death episodes for several days after exposure (Anderson et al., 1997; Katsouyanni et al., 1997; Schwartz, 2000b). Our 1 to 14-day estimates portray that the effect of a single-day exposure to $PM_{2.5}$ persists in infants for around two weeks before leveling off to zero (Figure 5). The mortality effect peaks again in neonates on the 10^{th} day, following the low estimate on day 7. Estimates for early neonates after the 6^{th} day are comparable to estimates for neonates, which conveys even a higher effect over their future neonatal period if ever reached. Presumably, the negative trend after the 4^{th} day indicates a harvesting effect (a period of fewer deaths following a period of more deaths, known as 'mortality displacement'), but this might be a short-term phenomenon only. Although pollution kills only vulnerable children up to 3-days (early neonates and neonates) and 4-days (infants and under five), more healthy children are added to the vulnerability pool in the period of fewer deaths. However, the fluctuating mortality pattern (Figure 5) depicts a possibility of future high-death episodes in children consonant with epidemiology literature (Schwartz, 2000a; Zanobetti, 2000), though our estimates do not predict the duration of the vulnerability pool.

A gold nanoparticles experiment, as a proxy for $PM_{2.5}$, confirms that particles move from lungs to blood and urine within 24 hours of exposure, stay in the bloodstream for at least 3 months, pile up in arteries, lead to blood clots, and affect the normal functioning of a human heart.[16] Thus, particulate matter can take weeks or months to exit the human body upon transient exposure, leaving the exposed person susceptible to further health shocks (Borgschulte et al., 2024). However, once children die, there is no way of clearing pollutants out of their system. For those who are still alive, the subsequent exposure may translate into a death incident, acting as a trigger for ongoing illness following previous transient or long-term exposures. While Deryugina et al. (2019) found a longer-window mortality effect beyond weeks of exposure for adults, a significant effect of up to 5 days in infants confirms their vulnerability compared to older adults.

6 Robustness check

We assume all census divisions in a cluster experience wind and, thus, pollution from the same direction as wind blows from a particular direction for each 10-degree wind bin. Therefore, the monotonicity assumption entails that $PM_{2.5}$ increase in the same direction for all CDs in a monitor group when the wind blows from a more polluted area to a less polluted area and vice versa (Angrist & Imbens, 1995; Deryugina et al., 2019). However, all CDs in a monitor group are less likely to face pollution from the same direction if CD-to-CD variation in landscapes changes the direction of the wind in each CD, if the direction of pollution varies across all the 36 wind direction bins, or if the relationship between $PM_{2.5}$ and wind direction is time-variant.

Though our specification employs the maximum possible wind bins, we estimate the same models by increasing the size of wind direction bins and lowering the number of clusters. We increase the size of wind direction bins from 10 degrees to 45 and 90 degrees in columns (1) and (2) and reduce the cluster size from 17 to 15 and 6 clusters in columns (3) and (4) in Table

[16] More from British Heart Foundation at https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2020/february/can-air-pollution-kill-you

4. All estimates from column (1) to column (4) are quite similar to our main specification in column (5) with 17 clusters and 10-degree wind bins, which validates the monotonicity assumption. Moreover, these results also support our idea of capturing non-local pollution for all monitors in a CD, as the first-stage estimates are reasonably good for all alternative wind bins and cluster choices.

Next, we estimate our chosen specification using different types of fixed effects and separate weather controls. Columns (1) to (5) in Table 5 show that our main results are robust with different fixed effects while including or excluding weather controls. Table 6 includes each weather control individually, and these results are consistent with Table 2. Therefore, Tables 5 and 6 suggest that topographical or meteorological conditions do not drive results in our preferred specification. However, our additional exercises show that estimates for neonates and children under 5 years are significantly high in the winter months and significantly low during the summer months, which supports the general U-shaped mortality pattern in the coldest countries (Falagas et al., 2009). Though $PM_{2.5}$ concentration is usually higher in summer than in winter; it is evident that the compound effect of pollution and influenza in winter increases virus-borne diseases and, hence, respiratory hospitalization rates (Graff Zivin et al., 2023).

We employ a placebo test assigning random pollution exposure (100–250 days before and after) to a death incident in Table 8 and a falsification test on injury-related death in Table 9. None of these estimates are significant, which supports the validity of our instrument again. Though our empirical approach includes numerous instruments and the first-stage F-statistics are very large, we re-estimate our preferred specification using a Limited Information Maximum Likelihood (LIML) estimator. The estimates in Table 10 are similar to the 2SLS estimates in Table 2, which reassures the strength of our instrument.

Figure 6 compares the non-linear effects of “Low” (< 3), “High” (6-9), and “Very High” (> 9) concentrations of $PM_{2.5}$ on child mortality relative to “Moderate” (3-6) concentration.[17] The coefficients for days with concentrations less than 3 μg/m³ are negative, but estimates for Neonates and Infants are significant on days $PM_{2.5}$ exceeding 9 μg/m³. The coefficient for infant mortality is also significant at a “High” level of $PM_{2.5}$. In general, this implies that our main results are driven by a $PM_{2.5}$ concentration greater than 6 μg/m³, which is approximately the mean concentration in our data.

Further, we re-estimate equation (1) clustering standard errors at various spatiotemporal levels, such as CD-month, province, province-year, year-month, and NAPS stations in Table 11. Results from alternative standard errors are slightly higher in magnitude and maintain the significance level in Table 2.

Additionally, we add the daily average of Ozone concentration with $PM_{2.5}$ to see how including more than one pollutant changes our preferred estimates. Table 12 shows that estimates for each death category are even higher than earlier results. Moreover, Ozone has a significantly negative effect on all death categories, which is directly comparable to Deryugina et al. (2019).

[17] Our qualitative categorization of daily average $PM_{2.5}$ readings is similar to the AQHI index published by Environment and Climate Change Canada (ECCC).
For details, please see https://weather.gc.ca/airquality/pages/index_e.html

7 Conclusion

Using the daily variation in ambient $PM_{2.5}$ driven by wind direction, we find a significant increase in all-cause child mortality in Canada. While most of the literature focuses on infant or under-five mortality rates, segregating age groups opens a new window to examine the exact period when they are most vulnerable to acute pollution exposure between zero and five years of age. To our knowledge, this is the first paper to analyze the effect of air pollution on child mortality at different stages of infanthood in children under five years of age. More specifically, considering child age in terms of days instead of years allows us to show that one-day exposure to $PM_{2.5}$ increases mortality rates in early neonates (0–6 days) and neonates (0–28 days) over 3 days. Furthermore, we show that the effect on male children is significantly higher than that on female children in all death categories, which establishes them as the most vulnerable child group. Our results support all policies to combat air pollution to save our future human capital. Likewise, we hope these insights will improve the adoption of personal protection and motivate possible avoidance for children in vulnerable age groups, creating more awareness among parents and caregivers.

References

Alnwisi, S. M. M., Chai, C., Acharya, B. K., Qian, A. M., Zhang, S., Zhang, Z., Vaughn, M. G., Xian, H., Wang, Q., & Lin, H. (2022). Empirical dynamic modeling of the association between ambient PM2.5 and under-five mortality across 2851 counties in Mainland China, 1999–2012. *Ecotoxicology and Environmental Safety*, *237*, 113513.

Anderson, H. R., Spix, C., Medina, S., Schouten, J. P., Castellsague, J., Rossi, G., Zmirou, D., Touloumi, G., Wojtyniak, B., & Ponka, A. (1997). Air pollution and daily admissions for chronic obstructive pulmonary disease in 6 European cities: results from the APHEA project. *European Respiratory Journal*, *10*(5), 1064–1071.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, *90*(430), 431–442.

Anwar, A., Ullah, I., Younis, M., & Flahault, A. (2021). Impact of air pollution (PM2.5) on child mortality: Evidence from sixteen Asian countries. *International Journal of Environmental Research and Public Health 18*(12).

Bachwenkizi, J., Liu, C., Meng, X., Zhang, L., Wang, W., van Donkelaar, A., Martin, R. V, Hammer, M. S., Chen, R., & Kan, H. (2021). Fine particulate matter constituents and infant mortality in Africa: A multicountry study. *Environment International*, *156*, 106739.

Bailey, M. J., Holzhausen, E. A., Morgan, Z. E. M., Naik, N., Shaffer, J. P., Liang, D., Chang, H. H., Sarnat, J., Sun, S., & Berger, P. K. (2022). Postnatal exposure to ambient air pollutants is associated with the composition of the infant gut microbiota at 6-months of age. *Gut Microbes*, *14*(1), 2105096.

Borgschulte, M., Molitor, D., & Zou, E. Y. (2024). Air pollution and the labor market: Evidence from wildfire smoke. *Review of Economics and Statistics*, *106* (6), 1558–1575.

Cai, J., Shen, Y., Zhao, Y., Meng, X., Niu, Y., Chen, R., Quan, G., Li, H., Groeger, J. A., & Du, W. (2023). Early-Life Exposure to PM2.5 and sleep disturbances in preschoolers from 551 cities of China. *American Journal of Respiratory and Critical Care Medicine*, *207*(5), 602–612.

Cakmak, S., Hebbern, C., Pinault, L., Lavigne, E., Vanos, J., Crouse, D. L., & Tjepkema, M. (2018). Associations between long-term PM2. 5 and ozone exposure and mortality in the Canadian Census Health and Environment Cohort (CANCHEC), by spatial synoptic classification zone. *Environment International*, *111*, 200–211.

Chen, H., Quick, M., Kaufman, J. S., Chen, C., Kwong, J. C., van Donkelaar, A., Meng, J., Martin, R. V, Kim, J., & Lavigne, E. (2022). Impact of lowering fine particulate matter from major emission sources on mortality in Canada: A nationwide causal analysis. *Proceedings of the National Academy of Sciences*, *119*(49), e2209490119.

Christidis, T., Erickson, A. C., Pappin, A. J., Crouse, D. L., Pinault, L. L., Weichenthal, S. A., Brook, J. R., van Donkelaar, A., Hystad, P., & Martin, R. V. (2019). Low concentrations of fine particle air pollution and mortality in the Canadian Community Health Survey cohort. *Environmental Health*, *18*, 1–16.

Chay, K. & Greenstone, M. (2003a). 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.

Chay, K., & Greenstone, M. (2003b). Air quality, infant mortality, and the Clean Air Act of 1970. Working Paper No. 10053. National Bureau of Economic Research.

Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., van Donkelaar, A., Martin, R. V, Villeneuve, P. J., Jerrett, M., Goldberg, M. S., & Pope III, C. A. (2015). Ambient PM2. 5, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environmental Health Perspectives*, *123*(11), 1180–1186.

Davila Cordova, J. E., Tapia Aguirre, V., Vasquez Apestegui, V., Ordoñez Ibarguen, L., Vu, B. N., Steenland, K., & Gonzales, G. F. (2020). Association of PM 2.5 concentration with health center outpatient visits for respiratory diseases of children under 5 years old in Lima, Peru. *Environmental Health*, *19*, 1–6.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178–4219.

Dunea, D., Iordache, S., Liu, H.-Y., Bøhler, T., Pohoata, A., & Radulescu, C. (2016). Quantifying the impact of PM 2.5 and associated heavy metals on respiratory health of children near metallurgical facilities. *Environmental Science and Pollution Research*, *23*, 15395–15406.

Egondi, T., Ettarh, R., Kyobutungi, C., Ng, N., & Rocklöv, J. (2018). Exposure to outdoor particles (PM2.5) and associated child morbidity and mortality in socially deprived neighborhoods of Nairobi, Kenya. *Atmosphere*, *9*(9), 351.

Elliott, C. T., & Copes, R. (2011). Burden of mortality due to ambient fine particulate air pollution (PM 2.5) in interior and Northern BC. *Canadian Journal of Public Health*, *102*, 390–393.

Falagas, M., Karageorgopoulos, D., Moraitis, L., Vouloumanou, E., Nikolaos, R., Peppas, G., & Rafailidis, P. (2009). Seasonality of mortality: The September phenomenon in Mediterranean countries. *CMAJ: Canadian Medical Association Journal*, *181*, 484–486.

Farhat, N., Ramsay, T., Jerrett, M., & Krewski, D. (2013). Short-Term effects of ozone and $PM_{2.5}$ on mortality in 12 Canadian Cities. *Journal of Environmental Protection*, *04*, 18–32.

Gisler, A., Korten, I., de Hoogh, K., Vienneau, D., Frey, U., Decrue, F., Gorlanova, O., Soti, A., Hilty, M., Latzin, P., & Usemann, J. (2021). Associations of air pollution and greenness with the nasal microbiota of healthy infants: A longitudinal study. *Environmental Research*, *202*, 111633.

Gogna, P., Narain, T. A., O'Sullivan, D. E., Villeneuve, P. J., Demers, P. A., Hystad, P., Brenner, D. R., Friedenreich, C. M., King, W. D., & Team, C. S. (2019). Estimates of the current and future burden of lung cancer attributable to PM2.5 in Canada. *Preventive Medicine*, *122*, 91–99.

Goyal, N., Karra, M., & Canning, D. (2019). Early-life exposure to ambient fine particulate air pollution and infant mortality: pooled evidence from 43 low-and middle-income countries. *International Journal of Epidemiology*, *48*(4), 1125–1141.

Graff Zivin, J., & Neidell, M. (2009). Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management*, *58*(2), 119–128.

Graff Zivin, J., Neidell, M., Sanders, N. J., & Singer, G. (2023). When externalities collide: Influenza and pollution. *American Economic Journal: Applied Economics*, *15*(2), 320–351.

Grammarly. (2025). Grammarly (Version 14.1232.0) [Desktop app]. https://www.grammarly.com/desktop/mac

Greenstone, M., & Chay, K. (2005). Does air quality matter? Evidence from housing markets. *Journal of Political Economy*, *113*, 376–424.

Hassan, N., & Murad, S. M. W. (2022). Does air pollution increase child mortality? Evidence from 58 developing countries. *Environmental Science and Pollution Research*, *29*(19), 28913–28932.

He, C., Liu, C., Chen, R., Meng, X., Wang, W., Ji, J., Kang, L., Liang, J., Li, X., Liu, Y., Yu, X., Zhu, J., Wang, Y., & Kan, H. (2022). Fine particulate matter air pollution and under-5 children mortality in China: a national time-stratified case-crossover study. *Environment International*, *159*, 107022.

Heft-Neal, S., Burney, J., Bendavid, E., Voss, K. K., & Burke, M. (2020). Dust pollution from the Sahara and African infant mortality. *Nature Sustainability*, *3*(10), 863–871.

Katsouyanni, K., Touloumi, G., Spix, C., Schwartz, J., Balducci, F., Medina, S., Rossi, G., Wojtyniak, B., Sunyer, J., Bacharova, L., Schouten, J., Ponka, A., & Anderson, H. (1997). Short-term effects of ambient Sulphur dioxide and Particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *BMJ (Clinical Research Ed.)*, *314*, 1658–1663.

Khadka, A., & Canning, D. (2022). Understanding the pathways from prenatal and post-birth PM2.5 exposure to infant death: An Observational Analysis Using US Vital Records (2011–2013). *International Journal of Environmental Research and Public Health 19*(1).

Künzli, N., Jerrett, M., Mack, W. J., Beckerman, B., LaBree, L., Gilliland, F., Thomas, D., Peters, J., & Hodis, H. N. (2005). Ambient air pollution and atherosclerosis in Los Angeles. *Environmental Health Perspectives*, *113*(2), 201–206.

Lavigne, É., Bélair, M.-A., Do, M. T., Stieb, D. M., Hystad, P., van Donkelaar, A., Martin, R. V, Crouse, D. L., Crighton, E., & Chen, H. (2017). Maternal exposure to ambient air pollution and risk of early childhood cancers: a population-based study in Ontario, Canada. *Environment International*, *100*, 139–147.

Lavigne, É., Talarico, R., van Donkelaar, A., Martin, R. V, Stieb, D. M., Crighton, E., Weichenthal, S., Smith-Doiron, M., Burnett, R. T., & Chen, H. (2021). Fine particulate matter concentration and composition and the incidence of childhood asthma. *Environment International*, *152*, 106486.

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of The 5th Berkeley Symposium on Mathematical Statistics and Probability/University of California Press*.

Mao, G., Nachman, R. M., Sun, Q., Zhang, X., Koehler, K., Chen, Z., Hong, X., Wang, G., Caruso, D., & Zong, G. (2017). Individual and joint effects of early-life ambient PM 2.5 exposure and maternal pre-pregnancy obesity on childhood overweight or obesity. *Environmental Health Perspectives*, *125*(6), 67005.

Matz, C. J., Egyed, M., Xi, G., Racine, J., Pavlovic, R., Rittmaster, R., Henderson, S. B., & Stieb, D. M. (2020). Health impact analysis of PM2.5 from wildfire smoke in Canada (2013–2015, 2017–2018). *Science of The Total Environment*, *725*, 138506.

Milani, G. P., Cafora, M., Favero, C., Luganini, A., Carugno, M., Lenzi, E., Pistocchi, A., Pinatel, E., Pariota, L., & Ferrari, L. (2022). PM2.5, PM10 and bronchiolitis severity: a cohort study. *Pediatric Allergy and Immunology*, *33*(10), e13853.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics*, *23*(6), 1209–1236.

Ng, C., Malig, B., Hasheminassab, S., Sioutas, C., Basu, R., & Ebisu, K. (2017). Source apportionment of fine particulate matter and risk of term low birth weight in California: exploring modification by region and maternal characteristics. *Science of The Total Environment*, *605*, 647–654.

Ortigoza, A., Gouveia, N., Kephart, J., Prado-Galbarro, F., Friche, A. A., Sanchez, B. N., & Diez-Roux, A. V. (2021). Association between ambient PM2.5 and under-5, infant, and child mortality in Latin America, 2010–15: a longitudinal analysis. *The Lancet Planetary Health*, *5*, S16.

Ostro, B., Roth, L., Malig, B., & Marty, M. (2009). The effects of fine particle components on respiratory hospital admissions in children. *Environmental Health Perspectives*, *117*(3), 475–480.

Owili, P. O., Lien, W.-H., Muga, M. A., & Lin, T.-H. (2017). The associations between types of ambient PM2. 5 and under-five and maternal mortality in Africa. *International Journal of Environmental Research and Public Health*, *14*(4), 359.

Pappin, A. J., Christidis, T., Pinault, L. L., Crouse, D. L., Brook, J. R., Erickson, A., Hystad, P., Li, C., Martin, R. V, & Meng, J. (2019). Examining the shape of the association between low levels of fine particulate matter and mortality across three cycles of the Canadian Census Health and Environment Cohort. *Environmental Health Perspectives*, *127*(10), 107008.

Pinault, L., Tjepkema, M., Crouse, D. L., Weichenthal, S., van Donkelaar, A., Martin, R. V, Brauer, M., Chen, H., & Burnett, R. T. (2016). Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. *Environmental Health*, *15*, 1–15.

Proietti, E., Röösli, M., Frey, U., & Latzin, P. (2013). Air pollution during pregnancy and neonatal outcome: a review. *Journal of Aerosol Medicine and Pulmonary Drug Delivery*, *26*(1), 9–23.

Ranjbaran, M., Mohammadi, R., Yaseri, M., Kamari, M., Habibelahi, A., & Yazdani, K. (2020). Effect of ambient air pollution and temperature on the risk of stillbirth: a distributed lag nonlinear time series analysis. *Journal of Environmental Health Science and Engineering*, *18*, 1289–1299.

Requia, W. J., Adams, M. D., & Koutrakis, P. (2017). Association of PM2. 5 with diabetes, asthma, and high blood pressure incidence in Canada: A spatiotemporal analysis of the impacts of the energy generation and fuel sales. *Science of The Total Environment*, *584*, 1077–1083.

Schwartz, J. (2000a). Harvesting and long-term exposure effects in the relation between air pollution and mortality. *American Journal of Epidemiology*, *151*(5), 440–448.

Schwartz, J. (2000b). The distributed lag between air pollution and daily deaths. *Epidemiology*, *11*(3), 320–326.

Sheridan, P., Ilango, S., Bruckner, T. A., Wang, Q., Basu, R., & Benmarhnia, T. (2019). Ambient fine particulate matter and preterm birth in California: Identification of critical exposure windows. *American Journal of Epidemiology*, *188*(9), 1608–1615.

Shin, H. H., Gogna, P., Maquiling, A., Parajuli, R. P., Haque, L., & Burr, B. (2021). Comparison of hospitalization and mortality associated with short-term exposure to ambient ozone and PM2.5 in Canada. *Chemosphere*, *265*, 128683.

Spears, D., Dey, S., Chowdhury, S., Scovronick, N., Vyas, S., & Apte, J. (2019). The association of early-life exposure to ambient PM 2.5 and later-childhood height-for-age in India: An observational study. *Environmental Health*, *18*, 1–10.

Stieb, D. M., Chen, L., Hystad, P., Beckerman, B. S., Jerrett, M., Tjepkema, M., Crouse, D. L., Omariba, D. W., Peters, P. A., & van Donkelaar, A. (2016). A national study of the association between traffic-related air pollution and adverse pregnancy outcomes in Canada, 1999–2008. *Environmental Research*, *148*, 513–526.

Stieb, D. M., Lavigne, E., Chen, L., Pinault, L., Gasparrini, A., & Tjepkema, M. (2019). Air pollution in the week prior to delivery and preterm birth in 24 Canadian cities: a time to event analysis. *Environmental Health*, *18*, 1–11.

Sun, Q., Wang, A., Jin, X., Natanzon, A., Duquaine, D., Brook, R. D., Aguinaldo, J.-G. S., Fayad, Z. A., Fuster, V., & Lippmann, M. (2005). Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. *Jama*, *294*(23), 3003–3010.

Tecer, L. H., Alagha, O., Karaca, F., Tuncel, G., & Eldes, N. (2008). Particulate matter (PM2. 5, PM10-2.5, and PM10) and children's hospital admissions for asthma and respiratory diseases: A bidirectional case-crossover study. *Journal of Toxicology and Environmental Health, Part A*, *71*(8), 512–520.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*(4), 267–276.

To, T., Zhu, J., Terebessy, E., Zhang, K., Fong, I., Pinault, L., Jerrett, M., Robichaud, A., Ménard, R., van Donkelaar, A., Martin, R. V, Hystad, P., Brook, J. R., Dell, S., & Stieb, D. (2021). Does exposure to air pollution increase the risk of acute care in young children with asthma? An Ontario, Canada study. *Environmental Research*, *199*, 111302.

Wagstaff, A., O'Donnell, O., Van Doorslaer, E., & Lindelow, M. (2007). *Analyzing health equity using household survey data: a guide to techniques and their implementation*. World Bank Publications.

Wang, H., Zhang, H., Li, J., Liao, J., Liu, J., Hu, C., Sun, X., Zheng, T., Xia, W., & Xu, S. (2022). Prenatal and early postnatal exposure to ambient particulate matter and early childhood neurodevelopment: A birth cohort study. *Environmental Research*, *210*, 112946.

Wang, Y., Li, Q., Guo, Y., Zhou, H., Wang, X., Wang, Q., Shen, H., Zhang, Y., Yan, D., & Zhang, Y. (2018). Association of long-term exposure to airborne particulate matter of 1 μm or less with preterm birth in China. *JAMA Pediatrics*, *172*(3), e174872–e174872.

Weichenthal, S., Crouse, D. L., Pinault, L., Godri-Pollitt, K., Lavigne, E., Evans, G., van Donkelaar, A., Martin, R. V, & Burnett, R. T. (2016). Oxidative burden of fine particulate air pollution and risk of cause-specific mortality in the Canadian Census Health and Environment Cohort (CanCHEC). *Environmental Research*, *146*, 92–99.

Wong, S. L., Coates, A. L., & To, T. (2016). Exposure to industrial air pollutant emissions and lung function in children: Canadian Health Measures Survey, 2007 to 2011. *Health Reports*, *27*(2).

Woodruff, T. J., Parker, J. D., & Schoendorf, K. C. (2006). Fine particulate matter (PM2.5) air pollution and selected causes of postneonatal infant mortality in California. *Environmental health perspectives*, *114*(5), 786-790.

Xia, X., & Yao, L. (2019). Spatio-temporal differences in health effect of ambient PM 2.5 pollution on acute respiratory infection between children and adults. *IEEE Access*, *7*, 25718–25726.

Yang, B.-Y., Bloom, M. S., Markevych, I., Qian, Z. M., Vaughn, M. G., Cummings-Vaughn, L. A., Li, S., Chen, G., Bowatte, G., & Perret, J. L. (2018). Exposure to ambient air pollution and blood lipids in adults: The 33 communities Chinese health study. *Environment International*, *119*, 485–492.

Yorifuji, T., Kashima, S., & Doi, H. (2016). Acute exposure to fine and coarse particulate matter and infant mortality in Tokyo, Japan (2002–2013). *Science of The Total Environment*, *551–552*, 66–72.

Zang, H., Cheng, H., Song, W., Yang, M., Han, P., Chen, C., & Ding, R. (2019). Ambient air pollution and the risk of stillbirth: a population-based prospective birth cohort study in the coastal area of China. *Environmental Science and Pollution Research*, *26*, 6717–6724.

Zanobetti, A. (2000). Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics*, *1*(3), 279–292. https://doi.org/10.1093/biostatistics/1.3.279

Zhao, Q., Kress, S., Markevych, I., Berdel, D., von Berg, A., Gappa, M., Koletzko, S., Bauer, C.-P., Schulz, H., & Standl, M. (2021). Air pollution during infancy and lung function development into adolescence: The GINIplus/LISA birth cohorts study. *Environment International*, *146*, 106195.

Zwozdziak, A., Sówka, I., Willak-Janc, E., Zwozdziak, J., Kwiecińska, K., & Balińska-Miśkiewicz, W. (2016). Influence of PM 10 and PM 2.5 on lung function parameters in healthy schoolchildren—a panel study. *Environmental Science and Pollution Research*, *23*, 23892–23901.

Figures and Tables

Figure 1: Difference in pollution exposure from local vs. non-local pollution source

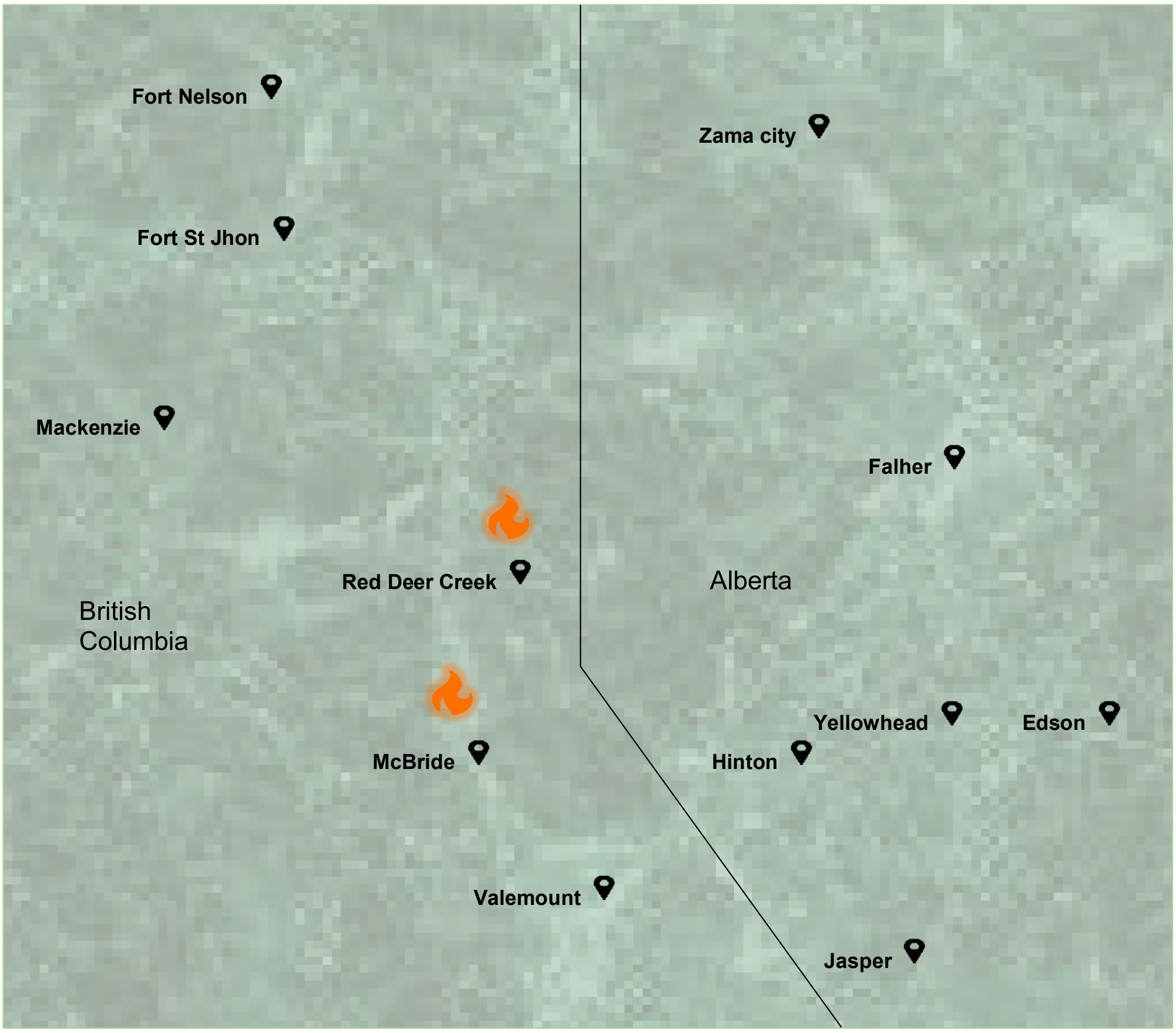
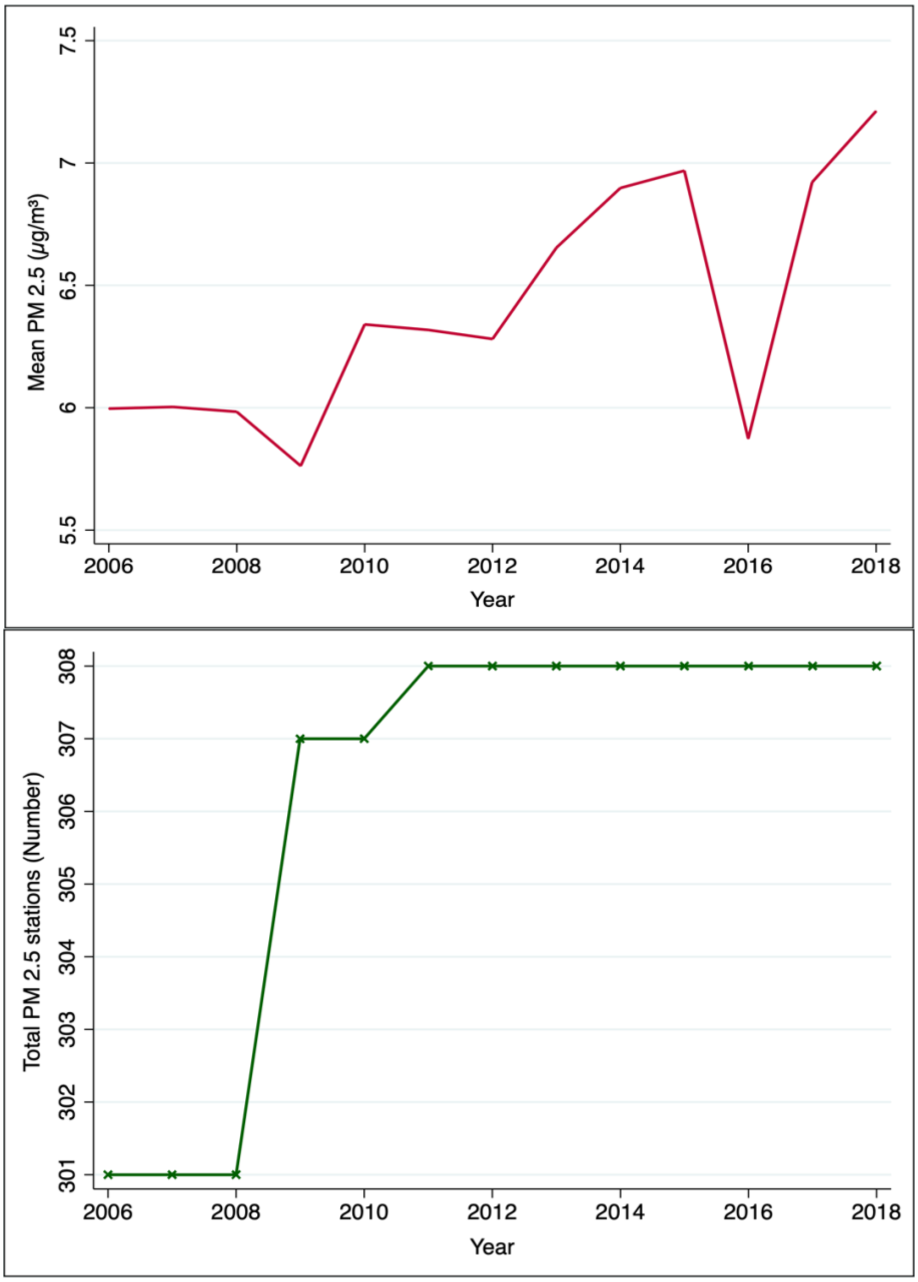
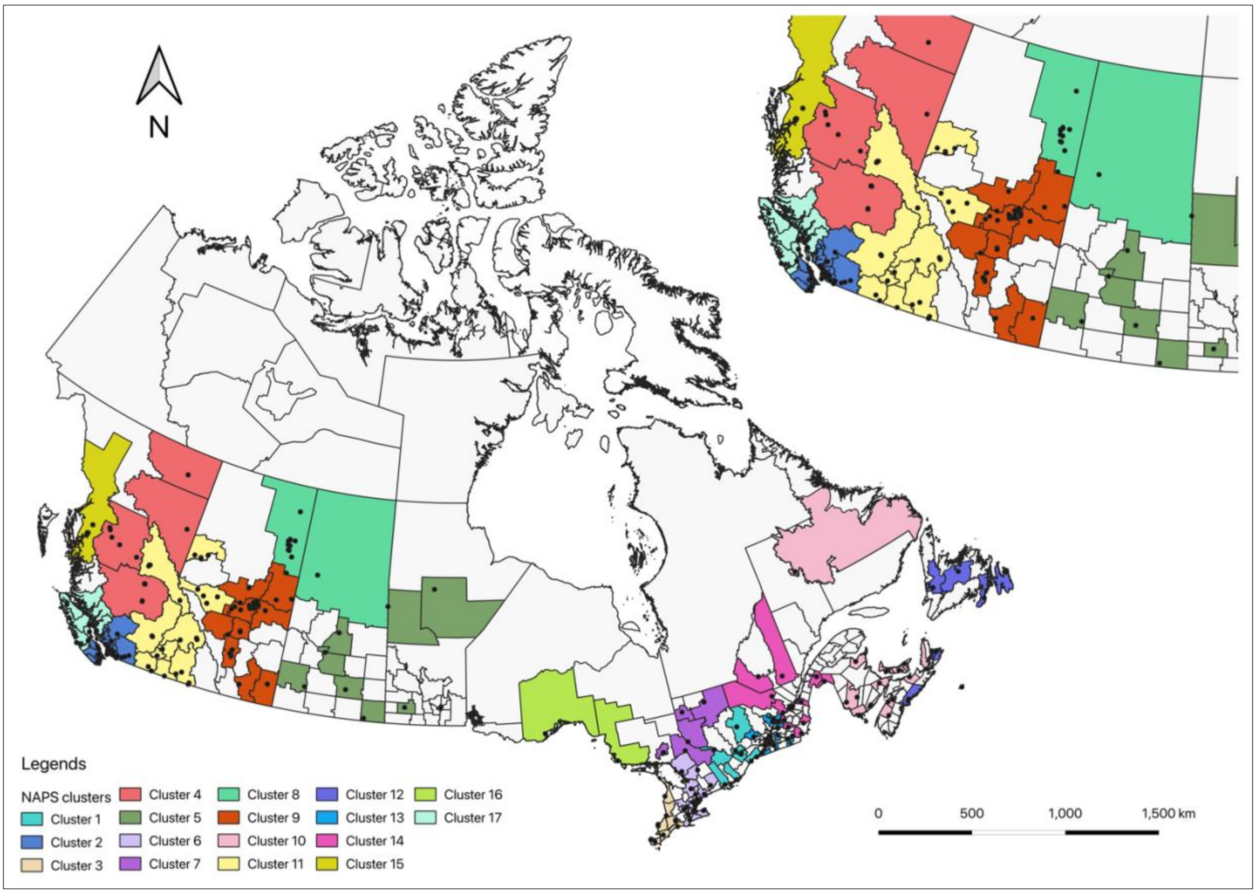


Figure 2: Trends in $PM_{2.5}$ concentration and NAPS stations from 2006 to 2018



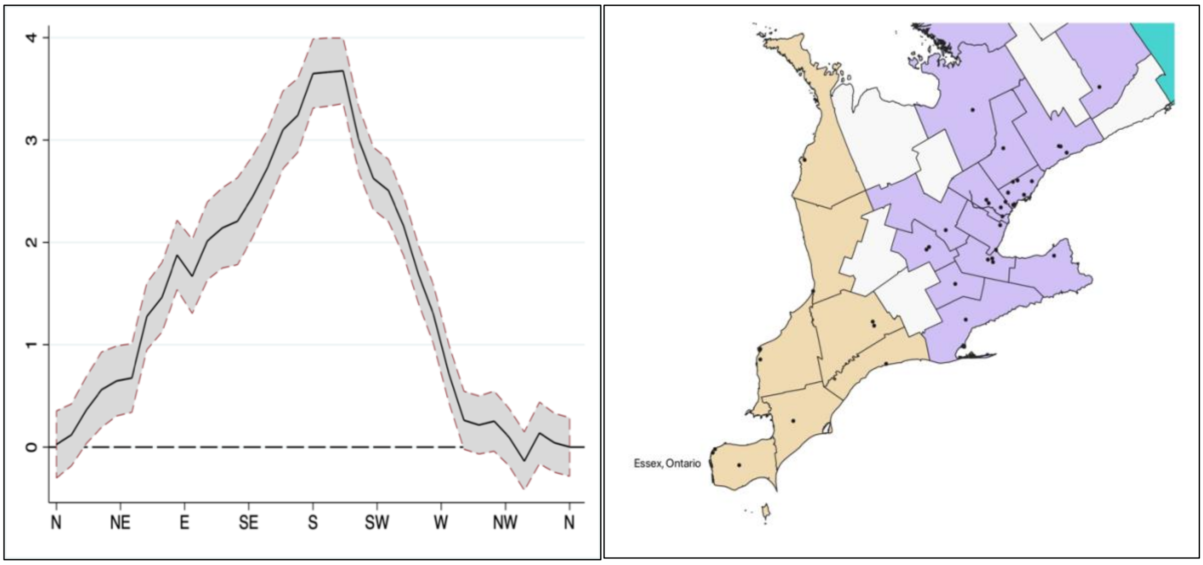
Notes: The upper panel exhibits the yearly average of $PM_{2.5}$ concentration across all census divisions, and the lower panel depicts the total number of NAPS stations in ten provinces.

Figure 3: Census divisions assigned to each NAPS cluster



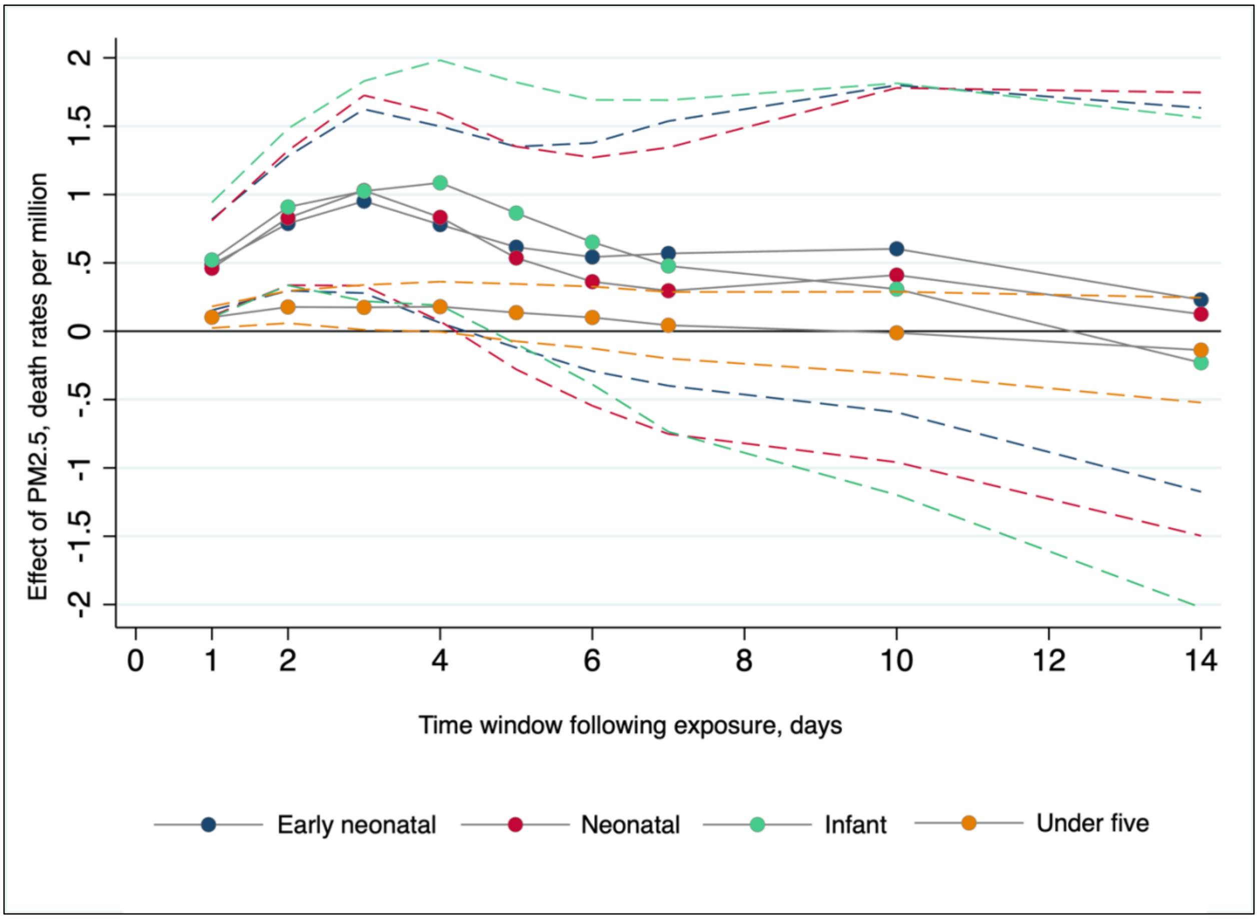
Notes: We allocate all census divisions into 17 NAPS clusters. Each legend represents a single cluster containing several CDs from one or more than one province. For example, cluster 4 includes CDs from British Columbia, whereas cluster 8 includes CDs from Alberta and Saskatchewan. Black dots represent the location of $PM_{2.5}$ stations. Due to the lack of stations, white CDs do not fall into any monitor group. The inset map provides a close-up view of the distribution of stations in British Columbia, Alberta, and Saskatchewan.

Figure 4: Regional wind direction and pollution in the Lake Erie region



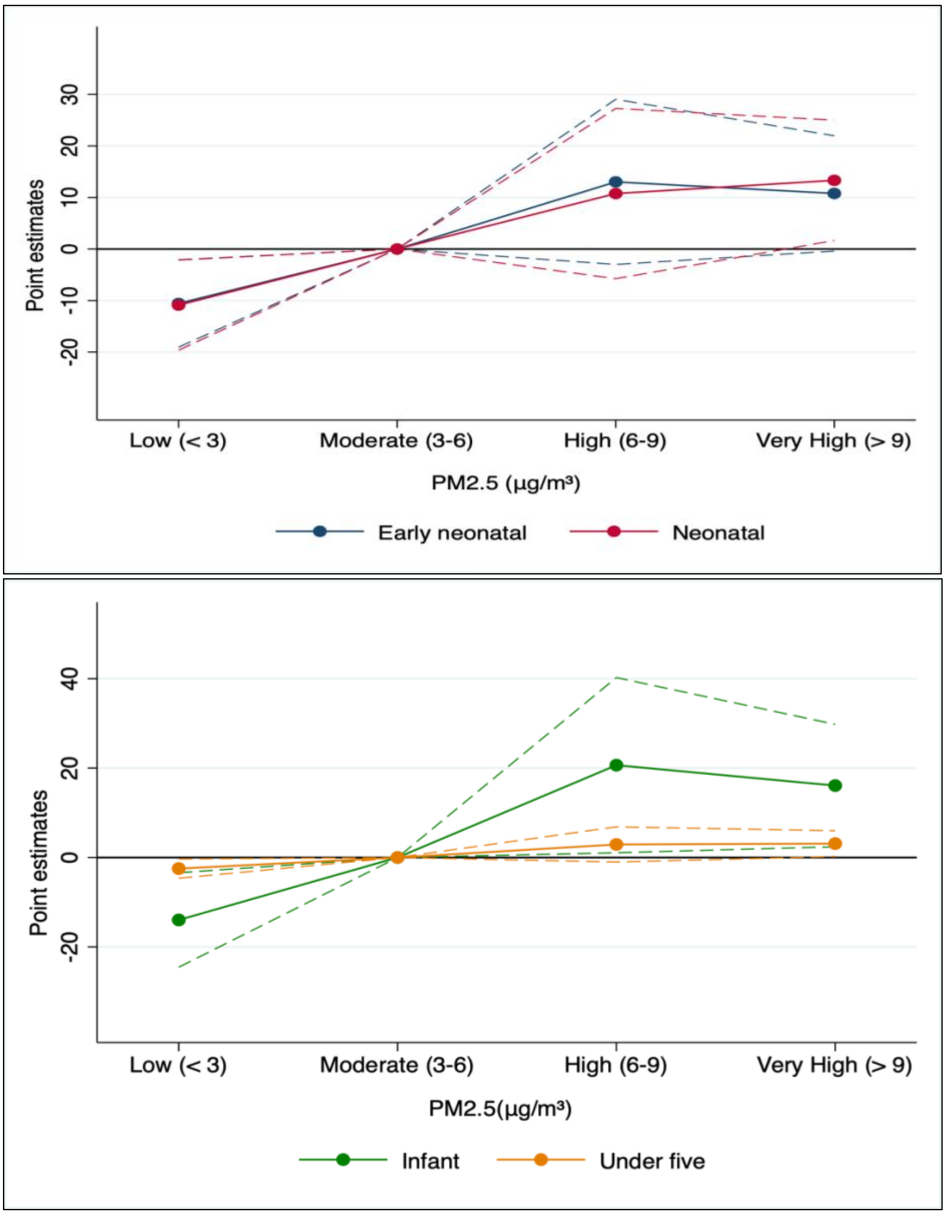
Notes: The left panel shows regression estimates from the first stage. The dependent variable is the average $PM_{2.5}$ concentration at the CD level. The regression includes an indicator variable for 10-degree wind bins, CD fixed effect, province-by-month fixed effect, month-by-year fixed effect, and weather variables from our first-stage specification in equation (1). The grey area inside the dashed line represents 95 percent confidence intervals based on robust standard errors. The black dots in the right panel depict the location of NAPS monitors in the Lake Erie region, providing pollution measures for this regression.

Figure 5: Mortality effect of a single day exposure to $PM_{2.5}$ over extended periods



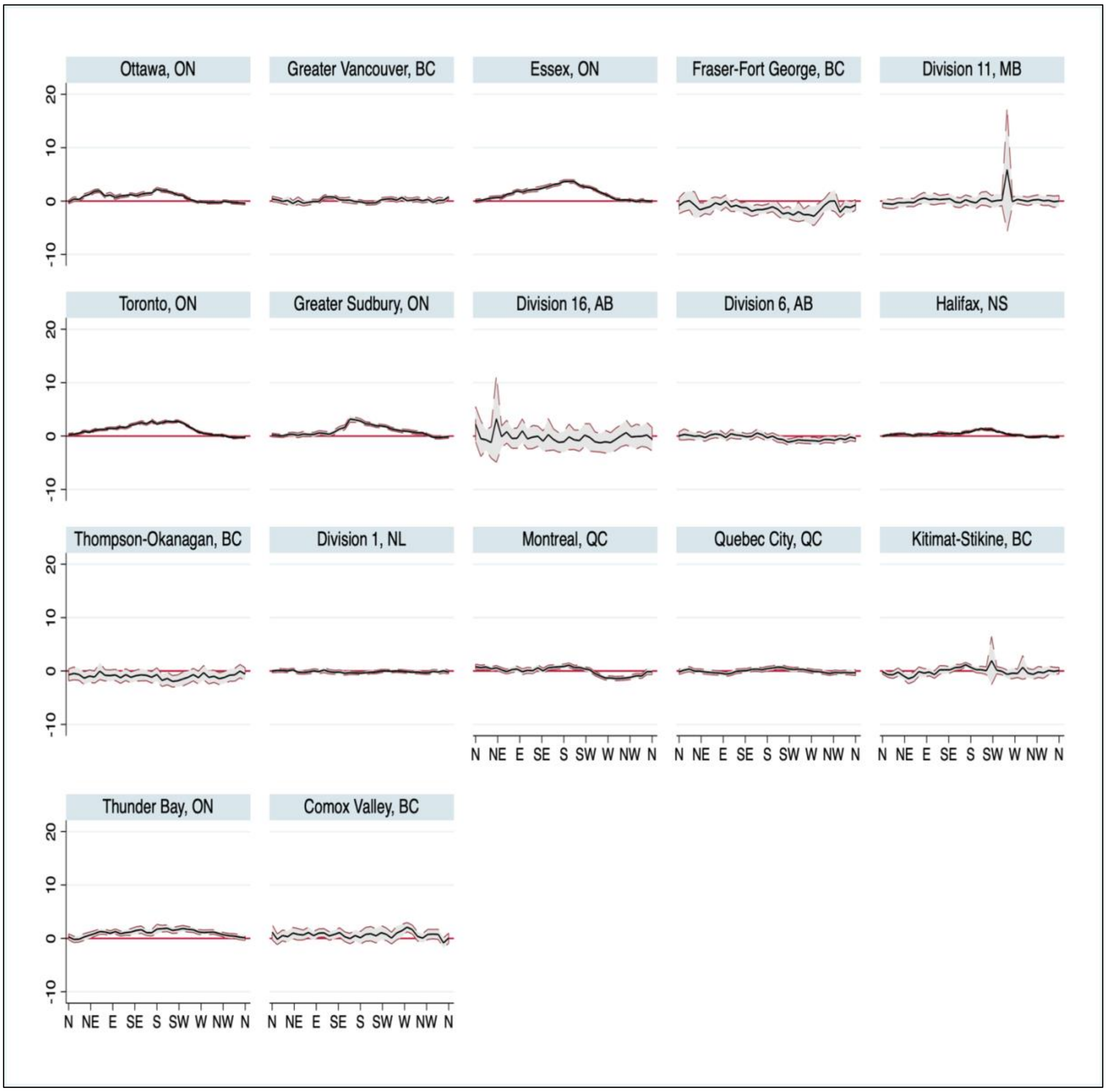
Notes: Figure 5 reports IV estimates of equation (2) up to 14 days following the exposure day. Outcome variables are mortality rate per million live births for early neonatal, neonatal, and infant, and per million children for under-five. Estimates show 1 day to 7 days, 10 days, and 14 days effect in all death categories. One-day and three-day estimates represent point estimates in Table 2 and Table 7. All regressions incorporate clusters and wind bins (instrument = clusters × wind bins), census division, province-by-month, and month-by-year fixed effects; an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover.

Figure 6: Non-linear estimates



Notes: Figure 6 reports IV estimates of equation (2) and corresponding 95% confidence intervals using binned values of daily average $PM_{2.5}$. Outcome variables are mortality rates per million live births for early neonatal, neonatal, and infant, and per million children for the under-five age category. All regressions incorporate the interaction of clusters and wind bins as instruments, census division, province-by-month, and month-by-year fixed effects; an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover.

Figure 7: First stage estimates by monitor groups



Notes: Figure 7 shows regression estimates from the first stage for 17 NAPS clusters. The dependent variable is the average $PM_{2.5}$ concentration at the CD level. The regression includes an indicator variable for 10-degree wind bins, CD fixed effect, province-by-month fixed effect, month-by-year fixed effect, and weather variables from our first-stage specification in equation (1). The grey area inside the dashed line represents 95 percent confidence intervals based on robust standard errors. We name each monitor group after the most densely populated region in that cluster.

Table 1: Summary statistics of sample data, 2006 – 2018

	Mean	Standard deviation	Observation
$PM_{2.5}$	6.440	10.069	500,310
Number of early neonatal death	0.021	0.158	500,310
Number of neonatal death	0.025	0.173	500,310
Number of post-neonatal death	0.007	0.084	500,310
Number of infant death	0.033	0.199	500,310
Number of one-five death	0.003	0.061	500,310
Number of under-five death	0.037	0.211	500,310
Number of early neonatal male death	0.011	0.113	500,310
Number of early neonatal female death	0.009	0.102	500,310
Number of neonatal male death	0.013	0.123	500,310
Number of neonatal female death	0.011	0.112	500,310
Number of post-neonatal male death	0.003	0.062	500,310
Number of post-neonatal female death	0.003	0.056	500,310
Number of infant male death	0.018	0.142	500,310
Number of infant female death	0.015	0.128	500,310
Number of one-five male death	0.002	0.044	500,310
Number of one-five female death	0.001	0.041	500,310
Number of under-five male death	0.020	0.150	500,310
Number of under-five female death	0.017	0.136	500,310
Three-day mortality rate, early neonatal death	22.518	193.910	500,310
Three-day mortality rate, neonatal death	27.116	211.361	500,310
Three-day mortality rate, post-neonatal death	9.428	127.706	500,310
Three-day mortality rate, infant death	36.771	247.368	500,310
Three-day mortality rate, one- five death	1.298	24.617	500,310
Three-day mortality rate, under- five death	8.335	52.227	500,310
Three-day mortality rate, early neonatal male death	23.641	273.167	500,310
Three-day mortality rate, early neonatal female death	21.354	268.370	500,310
Three-day mortality rate, neonatal male death	28.437	297.454	500,310
Three-day mortality rate, neonatal female death	25.736	293.956	500,310
Three-day mortality rate, post-neonatal male death	10.291	185.456	500,310
Three-day mortality rate, post-neonatal female death	8.518	176.200	500,310
Three-day mortality rate, infant male death	38.972	350.615	500,310
Three-day mortality rate, infant female death	34.460	344.071	500,310
Three-day mortality rate, one-five male death	1.397	37.154	500,310
Three-day mortality rate, one-five female death	1.184	30.756	500,310
Three-day mortality rate, under-five male death	8.855	74.449	500,310
Three-day mortality rate, under-five female death	7.776	71.609	500,310

Notes: Table 1 reports statistics of the estimation sample. The unit of observation is CD-date. One-five and under-five death rates are reported per million children in the relevant categories. All other death rates are reported as per million live births.

Table 2: OLS and IV effects of $PM_{2.5}$ on child mortality

	(1) Early neonatal	(2) Neonatal	(3) Post- neonatal	(4) Infant	(5) One-to-five	(6) Under five
			Panel A: OLS estimates			
Mean daily $PM_{2.5}$	-0.027	-0.029	-0.024	-0.053*	-0.001	-0.010*
	(0.019)	(0.022)	(0.014)	(0.030)	(0.002)	(0.006)
Observations	500,310	500,310	500,310	500,310	500,310	500,310
Dep. Var. mean	22.518	27.117	9.428	36.771	1.298	8.336
			Panel B: IV estimates			
Mean daily $PM_{2.5}$	0.952**	1.029**	0.020	1.025**	-0.034	0.174**
	(0.343)	(0.355)	(0.209)	(0.411)	(0.036)	(0.084)
Observations	500,310	500,310	500,310	500,310	500,310	500,310
F-statistic	16,295.48	16,295.48	16,295.48	16,295.48	16,295.48	16,295.48
Dep. Var. mean	22.518	27.117	9.428	36.771	1.298	8.336
Effect relative to mean, percent	4.228	3.795	0.212	2.788	-2.619	2.087

Notes: Table 2 reports OLS and IV estimates of equation (1). Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, post-neonatal, and infant, and per million children for one-to-five and under-five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover. OLS (IV) estimates include clusters and wind bins (instrument = clusters × wind bins). Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: IV effects of $PM_{2.5}$ by sex

	(1) Early neonatal	(2) Neonatal	(3) Infant	(4) Under five
Panel A: Male				
Mean daily $PM_{2.5}$	1.448**	1.459**	1.342**	0.215*
	(0.498)	(0.534)	(0.666)	(0.127)
Observations	500,310	500,310	500,310	500,310
F-statistic	16,295.48	16,295.48	16,295.48	16,295.48
Dep. Var. mean	23.641	28.437	38.972	8.855
Effect relative to mean, percent	6.125	5.131	3.443	2.428
Panel B: Female				
Mean daily $PM_{2.5}$	0.418	0.524	0.640	0.131
	(0.403)	(0.415)	(0.538)	(0.113)
Observations	500,310	500,310	500,310	500,310
F-statistic	16,295.48	16,295.48	16,295.48	16,295.48
Dep. Var. mean	21.354	25.736	34.461	7.777

Notes: Table 3 reports IV estimates of equation (1) by gender. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Robustness of IV effects on alternative instrument choices

	(1)	(2)	(3)	(4)	(5)
Wind bins (degrees)	90	45	10	10	10
Number of clusters	17	17	15	6	17
Panel A: Early neonatal					
Mean daily $PM_{2.5}$	0.881^{**}	0.643^{*}	0.893^{**}	0.931^{**}	0.952^{**}
	(0.384)	(0.355)	(0.337)	(0.343)	(0.343)
Panel B: Neonatal					
Mean daily $PM_{2.5}$	1.005^{**}	0.803^{**}	0.921^{**}	0.982^{**}	1.029^{**}
	(0.432)	(0.398)	(0.335)	(0.348)	(0.355)
Panel C: Infant					
Mean daily $PM_{2.5}$	1.248^{**}	0.983^{**}	0.931^{**}	0.917^{**}	1.025^{**}
	(0.510)	(0.459)	(0.404)	(0.412)	(0.411)
Panel D: Under five					
Mean daily $PM_{2.5}$	0.217^{**}	0.178^{*}	0.180^{**}	0.155^{*}	0.174^{**}
	(0.110)	(0.097)	(0.088)	(0.090)	(0.084)
Observations	500,310	500,310	500,310	500,310	500,310
F-statistic	54,846.93	281,371.68	536,000,000	2,471.26	16,295.48

Notes: Table 4 reports IV estimates of equation (1) using alternative instrument choices. The main specification in other tables considers 17 clusters and 10-degree wind bins. In this table, we also interact 45-degree and 90-degree wind bins with the preferred cluster and 6 and 15 clusters with the preferred 10-degree wind bins. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness of IV effects to alternative fixed effect choices and weather controls

Mean daily $PM_{2.5}$	(1)	(2)	(3)	(4)	(5)
Early neonatal	0.649^{**}	0.749^{**}	0.743^{**}	0.978^{**}	0.986^{**}
	(0.254)	(0.250)	(0.258)	(0.330)	(0.337)
Neonatal	0.742^{**}	0.849^{***}	0.841^{***}	1.028^{**}	1.042^{**}
	(0.251)	(0.248)	(0.255)	(0.337)	(0.343)
Infant	0.834^{**}	0.879^{**}	0.919^{**}	0.937^{**}	1.014^{**}
	(0.294)	(0.283)	(0.295)	(0.384)	(0.395)
Under five	0.154^{**}	0.162^{**}	0.171^{**}	0.161^{**}	0.178^{**}
	(0.060)	(0.059)	(0.062)	(0.080)	(0.082)
Weather controls	None	None	None	Full	Full
Census division FE	X	X	X	X	X
Month FE		X		X	
Year FE		X		X	
Year-by-month FE	X		X		X
Province-by-month FE	X				
Observations	507,930	507,930	507,930	500,310	500,310
F-statistic	253.19	4,816.35	544.63	35,000,000,000	2,108.02

Notes: Table 5 reports IV estimates of equation (1) with different fixed effects, including and excluding weather controls. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: IV effect with preferred FE and separate weather controls

Mean daily $PM_{2.5}$	(1) Temperature	(3) Wind speed	(4) Precipitation	(5) Humidity	(6) Cloud cover
Early neonatal	0.837**	0.677**	0.628**	0.661**	0.640**
	(0.294)	(0.272)	(0.248)	(0.254)	(0.253)
Neonatal	0.923**	0.767**	0.713**	0.757**	0.733**
	(0.311)	(0.267)	(0.245)	(0.250)	(0.251)
Infant	0.904**	0.856**	0.790**	0.850**	0.823**
	(0.358)	(0.307)	(0.286)	(0.294)	(0.294)
Under five	0.156**	0.153**	0.146**	0.161**	0.152**
	(0.073)	(0.063)	(0.059)	(0.061)	(0.060)
Observations	500,905	507,930	507,930	500,685	507,930
F-statistic	38,086.68	267.46	517.08	1,606.91	244.99

Notes: Table 6 reports IV estimates of equation (1) with different fixed effects, including weather controls separately. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: IV effects of $PM_{2.5}$ on single-day mortality rate

	(1) Early neonatal	(2) Neonatal	(3) Infant	(4) Under five
Mean daily $PM_{2.5}$	0.486**	0.460**	0.521**	0.102**
	(0.170)	(0.178)	(0.215)	(0.040)
Observations	500,520	500,520	500,520	500,520
F-statistic	8,611.40	8,611.40	8,611.40	8,611.40

Notes: Table 7 reports IV estimates of equation (1) with the outcome variables, the one-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Placebo test-100 to 250 days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure days	+100	-100	+150	-150	+200	-200	+250	-250
				Panel A: Early neonatal				
Mean daily $PM_{2.5}$	0.156	0.259	-0.192	0.538	-0.185	-0.471	-0.164	0.029
	(0.285)	(0.331)	(0.278)	(0.371)	(0.301)	(0.303)	(0.388)	(0.344)
				Panel B: Neonatal				
Mean daily $PM_{2.5}$	0.110	0.239	-0.139	0.419	-0.130	-0.488	-0.312	0.443
	(0.308)	(0.333)	(0.348)	(0.407)	(0.357)	(0.341)	(0.427)	(0.381)
				Panel C: Infant				
Mean daily $PM_{2.5}$	-0.037	0.079	-0.576	0.384	-0.243	-0.533	-0.649	0.583
	(0.379)	(0.341)	(0.421)	(0.487)	(0.397)	(0.386)	(0.461)	(0.419)
				Panel D: Under five				
Mean daily $PM_{2.5}$	0.005	0.018	-0.208**	0.024	-0.106	-0.062	-0.108	0.120
	(0.082)	(0.073)	(0.097)	(0.098)	(0.088)	(0.077)	(0.106)	(0.081)
Observations	489,176	491,495	483,660	486,925	478,065	482,300	472,525	477,570
F-statistic	668.17	373.32	935.36	556.85	1,0636.19	15,647.19	6,525.58	329.35

Notes: Table 8 reports IV estimates of equation (1). Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Falsification test- injury-related death events

	(1) All ages
Mean daily $PM_{2.5}$	-0.016
	(0.011)
Observations	500,320
F-statistic	9,113.72

Notes: Table 9 reports IV estimates of equation (1). Outcome variables are the three-day total mortality rate of injury-related deaths for all ages. Regression incorporates census division, province-by-month, and month-by-year fixed effects; an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover. Standard errors in parentheses are clustered at the CD level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 10: Alternative estimator- LIML effects of $PM_{2.5}$ on child mortality

	(1) Early neonatal	(2) Neonatal	(4) Infant	(6) Under five
Mean daily $PM_{2.5}$	1.136**	1.225**	1.219**	0.208**
	(0.414)	(0.428)	(0.492)	(0.101)
Observations	500,310	500,310	500,310	500,310
F-statistic	16,295.48	16,295.48	16,295.48	16,295.48
Dep. Var. mean	22.518	27.117	36.771	8.336
Effect relative to mean, percent	5.045	4.517	3.315	2.495

Notes: Table 10 reports IV estimates of equation (1) using the LIML estimator instead of the 2SLS estimator. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover. Standard errors in parentheses are clustered at the CD level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 11: Alternative standard errors

Mean daily $PM_{2.5}$	(1) CD-month	(3) Province	(4) Province-year	(5) Year-month	(6) NAPS station
Early neonatal	0.963^{**}	0.963^{**}	0.963^{**}	0.963^{**}	0.963^{**}
	(0.342)	(0.517)	(0.351)	(0.331)	(0.343)
Neonatal	1.054^{**}	1.054^{**}	1.054^{**}	1.054^{**}	1.054^{**}
	(0.365)	(0.596)	(0.379)	(0.351)	(0.353)
Infant	1.032^{**}	1.032^{**}	1.032^{**}	1.032^{**}	1.032^{**}
	(0.424)	(0.416)	(0.416)	(0.340)	(0.410)
Under five	0.174^{**}	0.174^{**}	0.174^{**}	0.174^{**}	0.174^{**}
	(0.088)	(0.058)	(0.082)	(0.084)	(0.084)
Observations	500,310	500,310	500,310	500,310	500,310

Notes: Table 11 reports IV estimates of equation (1) with alternative standard errors clustered at different spatiotemporal levels. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, cloud cover, and interaction of clusters and wind bins as an instrument. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Alternative pollutant- ozone

	(1) Early neonatal	(2) Neonatal	(4) Infant	(6) Under five
Mean daily $PM_{2.5}$	1.027^{**}	1.118^{**}	1.308^{**}	0.267^{**}
	(0.385)	(0.418)	(0.482)	(0.103)
Ozone	-0.666^{**}	-0.832^{**}	-0.845^{**}	-0.194^{**}
	(0.328)	(0.367)	(0.393)	(0.089)
Observations	462,555	462,555	462,555	462,555

Notes: Table 12 reports IV estimates of equation (1), including the daily average of ozone along with $PM_{2.5}$. Outcome variables are the three-day total mortality rate per million live births for early neonatal, neonatal, and infant, and per million children under five. All regressions incorporate census division, province-by-month, and month-by-year fixed effects, an indicator for temperature, a dummy for precipitation, wind speed, humidity, and cloud cover. Standard errors in parentheses are clustered at the CD level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 2 Effect of the 1997 Red River Flood on Children's Cognitive Ability

Abstract

This paper uses the historical 1997 Red River Flood in Manitoba as a natural experiment to investigate the causal impact of flood exposure on children's cognitive abilities. To establish this causality, we analyze children's standardized test scores and community-level flood exposure data, leveraging the latest continuous treatment approach of the Difference-in-Difference model. We find that a 1 km increase in the distance to the flood zone increases children's test scores by 0.4 points. Estimates show significant evidence of demographic influence, with sex and race being the most dominant factors. However, we do not find any influence of household income on this causal relationship. Our preferred results are consistent and withstand a series of robustness checks and falsification tests.

Keywords: 1997 Red River Flood – Flood exposure – Cognitive ability – Test score.

1 Introduction

The educational impacts of natural disasters are profound and multifaceted, affecting education systems (Wang, 2024; Esnard et al., 2018), students' learning experiences (Hebebci, 2023; Segarra-Alméstica et al., 2022; Onigbinde-Jacob, 2018), and the resources available to entire communities (Hussain & Mukhopadhyay, 2024; Eskander & Barbier, 2022). Rush, (2018) reports that natural disasters in Indonesia led to a decline in district-level student enrollment of 0.2% in primary schools and 1.7% in secondary schools. As a long-term consequence of disasters in Guatemala, Hermida, (2011) finds that children exposed to natural disasters in early childhood and during their school years, respectively, have 0.2 and 0.4 fewer years of schooling in adulthood.

A few studies examine the post-flood consequences on the education system and children's educational outcomes. Similarly, literature on the post-flood effects on standardized test scores in Canada is rare. We aim to fill these gaps in the literature by investigating the causal effect of exposure to the 1997 Red River flood on children's standardized test scores in Manitoba.

Flood events have both immediate and longer-term impacts on the exposed communities through several health and economic channels, such as evacuation and displacement (Munro et al., 2017), damaged homes (Graham & Rock, 2019), spread of infectious diseases (Shokri et al., 2020; Ashbolt, 2019), crop loss (Gould et al., 2020), injuries, death, and trauma (Tong, 2017). Recent pieces of literature investigate how flood alters human capital formation through its effect on children's educational outcomes, such as dropout and out-of-school status (Ahmed et al., 2022), loss of learning hours, high rate of absenteeism, and damaged school infrastructure (Lassa et al., 2023); Habiba et al., 2021; Chaudhary & Timsina, 2017; Akello, 2014; Mudavanhu, 2014). For instance, analyzing the Young Lives dataset developed by the University of Oxford Nguyen & Minh Pham (2018) shows that children from flood-affected households in Vietnam, Ethiopia, and

India do much worse in mathematics and vocabulary tests and are less likely to finish school and enroll in new classes. In Ethiopia, flood exposure reduces children's math and Peabody Picture Vocabulary Test (PPVT) scores by 9.3 percent and 2.9 percent of the mean scores, respectively. Moreover, the number of grade completions in children aged 12 to 15 years declines by 3.4 percent in Ethiopia, by 3.8 percent in India, and by 1.8 percent in Vietnam after children's exposure to flood. Lassa et al. (2023) report that the 2013 Jakarta flood affected at least 586 educational institutions, including 251 elementary and junior high schools, with five of these completely damaged. Saif-Ur-Rehman & Shaukat (2013) examine the effect of the 2010 flood in Pakistan on school damage in the Charsadda and Swat districts. They report that 23 schools in these two districts were damaged and remained closed for more than three months.

Another recent study by Mai & Hibiki (2023) discusses the short-term effect of flood exposure on school-going children's education, work-school balance, and cognitive scores in Vietnam. They found that immediate flood exposure reduces children's PPVT score on average by more than 0.3 standard deviation with a higher effect for the minority group. Moreover, they provide evidence of less school enrolment and a trade-off between study and work hours, implying a rise in child labor activities during the post-flood period in Vietnam. Thamtanajit (2020) provides further empirical evidence of the impact of a severe flood in 2011 in Thailand on national-level test scores. Using a difference-in-difference strategy, he finds that flood events affect student achievement by significantly lowering their high-stakes test scores in grades 6 and 9. In addition, flood exposure shows a negative association with grade progression, math, and reading skills for Indian female and marginalized students (Khalid et al., 2024). Although these papers provide an overall view of flood effects on children's cognitive abilities, such research does not exist in the context of Canada or North America.

1.1 Contribution

We examine the effect of the 1997 Red River flood on children's standardized test scores in Manitoba. Although our findings are not generalizable to the entire country, this is the first study to explore the causal effect of flood exposure on children's cognitive ability in Canada, and indeed, in North America. Despite the increasing frequency of flood events across North America, we have not found any studies analyzing the causal effect of flood exposure on children's test scores. Our study contributes to the disaster literature by adding evidence from a developed country perspective. This is important because our findings show that even high-income countries with better infrastructure, resources, and emergency response plans are still vulnerable to extreme weather events like floods.

Previous studies have examined post-flood health and economic outcomes, and the indirect effect on children's cognitive ability through damaged school infrastructure, school closures, dropouts, and high absence rates. Only a few papers establish causality between natural disasters including floods and standardized or national-level test scores for children in developing countries, focusing primarily on South Asia. For example, Khalid et al. (2024) use a fixed-effect regression and report that village-level flood exposure in India reduces children's math and reading skills by 0.1 standard deviations, respectively. Shidiqi et al. (2023) examine the causal effect of the 2006 earthquake on human capital formation as a natural experiment, finding that exposure to the earthquake during school years reduces years of schooling by about 0.7 years and decreases the likelihood of completing compulsory education by 10 to 11 percentage points, with larger impacts on younger children and children of less-educated mothers. In another study, Deuchert & Felfe (2015) examine the short- and long-term effects of the 1990 typhoon on children's education and health in the Philippines, finding that grade retention increases by about 0.5 years and IQ scores decrease by 0.1 over time.

Our study is closely related to Pelli & Tschopp (2025), who provide causal evidence on the effect of storms on early education and human capital in India. Using a cross-sectional cohort study, they find that storm exposure increases educational delays by 2.4%, decreases post-secondary education achievement, and reduces regular salaried job attainment by 2% and 1.6%, respectively. We add another piece of causal evidence showing that flood exposure significantly reduces children's test scores during infancy and school age in a developed country. This has important policy implications, given the long-term impacts on Canada's human capital resources and labor market efficiency.

Methodologically, our paper is most similar to Thamtanajit (2020), who employs a traditional difference-in-differences design to examine the 2011 flood's impact on national test scores in Thailand, finding significant reductions for grades 6 and 9 (ranging from 0.03 to 0.11) but no effect for grade 12. We contribute to methodological advancement in disaster research by introducing a continuous treatment measure for flood exposure, following Callaway et al. (2024), instead of the conventional binary before/after and treatment/ control design. The rationale is that extreme weather events can generate psychological stress even for those living far from the directly affected area. Therefore, the distance to the flood serves as a more nuanced treatment variable, capturing both direct and indirect impacts that may extend well beyond the flood's immediate boundaries.

2 Flood of 1997 in the Red River Valley

The flood of 1997 in the Red River Valley of Manitoba is one of Canada's historical floods and the most significant flood of the twentieth century. A record discharge along the river, driven by an unexpectedly heavy snowstorm in May 1997, inundated 1,840 km^2 of land in Manitoba, creating a water body 100 km long with a maximum width of about 40 km (Rannie, 2016). The flood forced 27,400 people to evacuate and caused damage of approximately $1 billion in direct and indirect costs. Pre-1997 ring-diked communities survived with minor damage after an emergency elevation of 0.9 to 1.4 meters above the peak flood point, except for Ste. Agathe and Grande Point. Evacuees in affected communities could not return home until several weeks after the crests.[18] As a measure of flood protection and floodproofing, ring dikes were constructed around 10 additional communities along with a new elevation standard of 0.6 meters above the 1997 level (Rannie, 2016).

3 Data

3.1 National Longitudinal Survey of Children and Youth (NLSCY)

NLSCY is a nationally representative longitudinal survey of Canadian children and youth. The survey is conducted every two years, jointly by Statistics Canada and Human Resources and Skills Development Canada (HRSDC). NLSCY covers various topics, such as education, learning, health, family functioning, custody, friends, activities, mental health, and behavior to measure children's developmental experiences. Respondent households from the Labour Force Survey (LFS) serve as the initial source of sample selection for NLSCY. Children and youths aged 0 to

---

[18] Details at https://www.gov.mb.ca/flooding/

11 years from non-institutionalized civilian families comprise the sample population for the first cycle conducted in 1994. The survey ended in 2008 with a sample coverage aged 0 to 25 years, adding older youths in subsequent cycles. As a measure of cognitive ability, NLSCY provides standard scores of three different tests: a motor and social development test for children aged below 4 years, Peabody Picture Vocabulary Tests (PPVT) for kids aged 4 and 5 years, and a math test for children in Grade 2 and above. In this analysis, we include children residing only in Manitoba from cycle 1 to cycle 8 to measure the effect of flooding on children's cognitive performance.

3.2 Measure of flood exposure

About 80% of urban cities in Canada are located entirely or partially in flood zones, leaving most city dwellers vulnerable to flooding.[19] The 1997 flood inundated an area of 1,840 km^2, which is about 0.3% of Manitoba's total land area of 650,000 km^2. All schools and highways in southern Manitoba were closed due to the flood.

Figure 8 panel (a) exhibits Manitoba's flood polygon and Forward Sortation Areas (FSAs) based on the 1996 census boundary files.[20] To assign flood exposure to individuals, we measure the distance from each FSA centroid to the flood boundary. Consequently, the distance variable represents flood exposure for all individuals within a specific FSA. The NLSCY survey period includes three census years: 1996, 2001, and 2006. We compute and compare distances across all three census boundary files, opting to use the distance of the 1996 boundary data because the

---

[19] Details at https://www.canada.ca/en/services/environment/weather/climatechange/climate-plan/national-adaptation-strategy/action-plan.html

[20] Cartographic boundary file of FSAs is available from Statistics Canada.
Details at https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2001-eng.cfm
The overland flooding boundary in the RED River Valley is available from the Manitoba Government's website. Details at https://geoportal.gov.mb.ca/datasets/manitoba::red-river-flood-1997/explore?location=49.410486%2C-97.351521%2C10.77

distance variation among the three files is minimal. Half of the 63 FSAs in the 1996 boundary file are within 20 kilometers of the floodwater, with a minimum distance of zero and a maximum of 740 kilometers. Two FSAs, R3V and R5A, were entirely submerged in flood water.

4 DID approach with continuous treatment

We aim to estimate the causal impact of flood exposure on children's test scores in Manitoba. We assign flood exposure to individuals at the FSA level and refer to it as community *c*. The key assumption for establishing causality is that if flood exposure affects children's cognitive abilities, then test scores for children living close to the flood boundary would decrease significantly after the flood compared to those living farther away.

Hence, we employ a generalized difference-in-differences (DID) method, controlling for potential confounders, such as children's age, sex, health condition, family income, and accounting for time-varying shocks, such as education reform, changes in school funding and infrastructure, and student-teacher ratio. We start with a continuous treatment distance to the flood polygon ($d \in D$), and by considering two periods, pre-flood ($f_b = 0$) and post-flood ($f_a = 1$), following the recent DID approach by Callaway et al. (2024).[21]

However, the continuous treatment in our setting forces us to think that everyone is treated in Manitoba in period F_a. This may leave our readers in wonder as we fall short of having a post-treatment comparison group to portray the parallel trend assumption. Therefore, we utilize a cut-

[21] Using a standard traditional two-period (before and after flood) and two-group (treatment and control) Difference-in-Difference setup requires considering respondents from the inundated communities during the flood as the treatment group and the rest of Manitoba as the control group. However, this can lead to a misleading counterfactual statement and unrealistic inference arising from two issues: i) The treated groups could be considered a comparison group. For example, individuals from FSAs far north of Manitoba can still experience the effects of floods. ii) There is a risk of negative weights on parameters for the treatment groups, as weights are driven by the estimation method. For instance, two-way fixed-effect estimates might suffer from negative weight issues due to differences in group size, average dose, or treatment timing for each group.

off radius of 100 kilometers to define our treatment group T in period F_a. Next, our hypothesis is that distance to flood (proximity to flood) produces a heterogeneous treatment effect to group T (consisting of 49 FSAs) for different distance values d, considering individuals living more than 100 km away as the control group C (composed of 14 FSAs).[22]

Our identifying assumption is that the effect on the test score of individual i with a slight marginal increase in the distance to flood provides a good counterfactual for the effect on the test score that would have been observed for individual j with a large marginal increase in the distance to flood if individual j's distance had changed similarly. In other words, individual i and individual j would have experienced an identical effect on the outcome variable if they had the same proximity to the flood. More precisely, we estimate equation (1) to determine the causal effect, if there is any, of flood exposure on children's standard test scores:

$$Y_{ict} = \beta(D_{ict} \times F_{ict}^{100km \cdot f_a}) + \chi_{ict}\varphi + \rho_c + \tau_t + \varepsilon_{ict} \tag{1}$$

where Y_{ict} is the test score of individual i of community c in year t. The interaction term $(D_{ict} \times F_{ict}^{100km \cdot f_a})$ includes a continuous measure and a factor notation of treatment. Individual i's distance to flood is D_{ict}; $F_{ict}^{100km \cdot f_a}$ implies 1 for individuals living within 100 km of the flooded area during the post-flood periods f_a.[23] Our parameter of interest β presents the effect of flood exposure or average causal response on the treated group (ACRT). More specifically, β captures the heterogeneous causal effect where a marginal increase in the proximity to flood (marginal decrease in the distance to flood) increases (decreases) the effect on the response variable.

[22] We aim to relax the *strong parallel trend assumption* and employ a modified identification assumption in our setting. Using a cut-off radius facilitates visualization of the standard parallel trend. See Callaway et al. (2024) for details.
[23] Cycle 3 (1998) to Cycle 8 (2008) in NLSCY.

Our more specific identifying assumption for ACRT is that starting with any particular distance d, individuals closer to the flood polygon would experience a higher treatment effect than individuals further away. To be more specific, it implies that, for all $d \in D[0, 100]$

$$\mathbb{E}[Y_{t=2}(d) - Y_{t=1}(0) \mid D = d\,] \geq \mathbb{E}[Y_{t=2}(d) - Y_{t=1}(0) \mid D = d + i,\ i \in (0, D_{max}^{100km}]\,] \quad (2)$$

Narratively, we have twofold assumptions here. First, the test scores of individuals who stay in the same proximity to the flood are comparable without selection bias since they are similar enough after controlling for their socio-demographic and economic attributes. Second, test scores of individuals with substantially different distances from the flood polygon are also comparable without any selection bias, as their location choice is not an informed decision considering the flood of 1997.

Next, we add the vector χ_{ict} to control for individual and household attributes, which include the child's age, sex, and health condition, parents' education level, and household income. We add a community fixed effect ρ_c and a year fixed effect τ_t to remove the influence of unobserved spatiotemporal factors. All regressions include standard errors clustered at the community level.

4.1 Identification

The primary identification assumption in this paper is that the 1997 Red River Flood was an exogenous shock caused by an unusually heavy snowfall in late spring. The event was unforeseen, and variation in exposure arose from individuals' random geographic proximity to the flood zone. Floods result from complex interactions of geological and atmospheric processes, including solar heating, rapid snowmelt, deforestation, and land use changes (Whitfield, 2012), which makes precise prediction difficult. Historical flood patterns in the Red River Valley did not provide reliable information about the likelihood or severity of future floods, and the effectiveness of preparedness measures depends on the unpredictable magnitude of each event.

Although Manitobans were aware of Red River's location and past floods, the presence of ring dikes prior to 1997 reduced the perceived likelihood of large-scale flooding and therefore did not drive residential location choices. Our measure of exposure—distance from each forward sortation area (FSA) centroid to the flood polygon—is effectively predetermined with respect to the 1997 event. Because individuals could not adjust their location choices in anticipation of the flood, and large-scale post-flood relocation did not occur, this distance measure is plausibly exogenous after controlling for community fixed effects.

Figure 8, panel (b), illustrates treatment and control areas. Treated FSAs are defined as those with centroids within 100 km of the floodplain (white areas), while control FSAs are those with centroids more than 100 km away (blue areas). These control FSAs are similar to treated FSAs in socio-economic characteristics, education systems, and climate conditions, providing a reasonable counterfactual for estimating the flood's causal impacts.

The identifying assumption for β is twofold:

1. Conditional on covariates (child's age, sex, health, parents' education, household income), fixed effects, and year effects, FSAs at different distances from the flood would have had parallel trends in test scores absent the flood.
2. Distance to the flood is not related to unobserved determinants of test scores because residential location was not chosen in anticipation of the 1997 event.

Our identification assumptions are likely to hold for three main reasons. First, the timing and severity of the flood were driven by extreme and atypical weather conditions, not by local economic or social factors that could also affect educational outcomes. Second, pre-flood infrastructure and historical experience with smaller floods created a low perceived risk, meaning settlement patterns were not sorted on anticipated flood exposure. Third, administrative and census records show no evidence of significant population composition changes between treated and control FSAs in the years immediately before or after the flood, limiting the risk of endogenous migration bias. Together, these factors support the parallel trends assumption and the plausibility of distance to the flood as an exogenous measure of treatment intensity.

A limitation is that we cannot present a formal visual parallel trend check, as our dataset contains only one year of pre-flood test score observations. However, the combination of the flood's unexpected nature, the absence of selective migration, and the similarity of treated and control FSAs in pre-flood characteristics supports the plausibility of the parallel trend assumption. Together, these factors justify using distance to the flood as an exogenous measure of treatment intensity.

5 Main Results

Based on the methods outlined in section 4, we estimate the causal effect of flood exposure on children's test scores. Column (1) in Table 14 shows the effect for all children, while columns (2) and (3) present it separately for male and female children. The fixed effect estimators from our difference-in-difference analysis indicate that proximity to flood adversely affects test scores for all children. We find that the average test score increases by approximately 0.4 points for every 1 km increase in distance from the flood. The effect on female children's test scores is 0.1 points higher than that for male children, suggesting that their cognitive abilities are more vulnerable to flood exposure.

In Table 15 and Table 16, we estimate the effect across three different age groups and races of children. Columns (1) to (3) in Table 15 indicate that flood exposure affects children's test scores in all age groups, and the effect size gradually increases for older children. In Table 16, we find a significant effect of flood exposure on the test scores of non-white children, whereas white children do not have any effect on their test scores.

This section presents additional subsample exercises considering children's prevailing health conditions, such as learning disability, birth prematurity, and long-term health issues, and household income status. We have found a significant effect of flood exposure on test scores for children without learning disability in Table 17. The sample size for disabled children is unexpectedly very low, but the estimate on disability interacted with the treatment in column (3) is statistically significant, which implies that the flood makes kids with disability more vulnerable to scoring low due to the exposure to the 1997 flood. However, our additional exercises do not show any effect of birth prematurity, long-term health conditions, or household income on test scores.

Furthermore, we have added additional exercises considering alternative radii of the flood-affected area, 50 km, 250 km, 500 km, and 750 km. We also conducted a falsification test considering all individuals beyond 500 km as flood-affected, i.e., considering the controls as treatment. Table 18 summarizes the results of additional exercises along with our preferred specification. The estimate with a radius of 50 km is close to our main result in magnitude, but the effect size decreases gradually for a 250 km radius and a 500 km radius. Considering a further radius of 750 km shows no effect of flood on test scores. However, we still find a significant estimate up to 500 km distance, which validates our identification with the continuous treatment approach. Moreover, the result of the falsification test in column (4) validates our initial hypothesis and overall findings.

6 Conclusion

Our preliminary results confirm that flood exposure harms children's cognitive abilities by revealing the significant adverse effects of the 1997 Red River flood on standard test scores in Manitoba. Using a generalized Difference-in-Difference method, we show that test scores decline by 0.4 points as proximity to the flood increases by 1 km for children within 100 km of the flood zone. Moreover, we found that the cognitive abilities of non-white children and children of all age groups are vulnerable to flood exposure. We find no evidence of a mitigating effect of household income or an augmenting effect of prevailing health conditions, except for learning disability. Our findings align with existing literature Mai & Hibiki (2023), though limited research in this area and differences in estimation methods make direct comparisons challenging. Furthermore, our limitation of data access and insufficient observations demand a further investigation of this research question.

References

Ahmed, R., Barkat, W., Ahmed, A., Tahir, M., & Nasir, A. M. (2022). The Impact of flooding on education of children and adolescents: Evidence from Pakistan. *Water Economics and Policy*, *08*(03), 2240009.

Akello, S. (2014). *Effects of floods on student's access to secondary education in Nyando District, Kisumu County, Kenya*. University of Nairobi.

Ashbolt, N. J. (2019). Flood and infectious disease risk assessment. *Health in Ecological Perspectives in the Anthropocene*, 145–159.

Callaway, B., Goodman-Bacon, A., & Sant'anna, P. H. C. (2024). Difference-in-differences with a continuous treatment, (NBER Working Papers) 32117, National Bureau of Economic Research, Inc.

Chaudhary, G., & Timsina, T. P. (2017). Impact of flood on performance of students the case of secondary school students in Jaleshwor Municipality, Mahottari. *Journal of Advanced Academic Research*, *4*(1), 12–26.

Deuchert, E. and Felfe, C. (2015). The tempest: short- and long-term consequences of a natural disaster for children's development. European Economic Review, 80, 280–294.

Eskander, S., & Barbier, E. (2022). Long-term impacts of the 1970 cyclone in Bangladesh. *World Development*, *152*, 105793.

Esnard, A.-M., Lai, B. S., Wyczalkowski, C., Malmin, N., & Shah, H. J. (2018). School vulnerability to disaster: examination of school closure, demographic, and exposure factors in Hurricane Ike's wind swath. *Natural Hazards*, *90*, 513–535.

Gould, I. J., Wright, I., Collison, M., Ruto, E., Bosworth, G., & Pearson, S. (2020). The impact of coastal flooding on agriculture: A case-study of Lincolnshire, United Kingdom. *Land Degradation & Development*, *31*(12), 1545–1559.

Graham, T. M., & Rock, M. J. (2019). The spillover effect of a flood on pets and their people: implications for rental housing. *Journal of Applied Animal Welfare Science*, *22*(3), 229–239.

Grammarly. (2025). Grammarly (Version 14.1232.0) [Desktop app]. https://www.grammarly.com/desktop/mac

Habiba, U., Jui, F. T., Meem, T., & Kabir, F. (2021). Disaster and education: Impact of flood on school going children at Keshabpur Upazilla in Bangladesh. *Journal of Disaster and Emergency Research*, *4*(2), 81-92.

Hebebci, M. (2023). *The impact of natural disasters on education*. International Conference on Social Science Studies.

Hermida, P. (2011). *The long-term effect of natural disasters: Health and education in Guatemala after the 1976 earthquake*. Social Science Research Network.

Heyes, A., & Saberian, S. (2024). Pollution and learning: causal evidence from Obama's Iran sanctions. *Journal of Environmental Economics and Management*, *125*, 102965.

Hussain, Y. R., & Mukhopadhyay, P. (2024). Impact of natural disasters on educational attainment in India: a panel data analysis. *Discover Sustainability*, *5*(1), 279.

Khalid, N., Behrman, J., Hannum, E., & Thapa, A. (2024). Floods and children's education in rural India. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *10*(1), 230–254.

Lassa, J., Petal, M., & Surjan, A. (2023). Understanding the impacts of floods on learning quality, school facilities, and educational recovery in Indonesia. *Disasters*, *47*(2), 412–436.

Mai, T. T. C., & Hibiki, A. (2023). *How does flood affect children differently? The impact of flood on children's education, labor, food Consumption, and cognitive development*. TUPD Discussion Papers, 36, Graduate School of Economics and Management, Tohoku University.

Mudavanhu, C. (2014). The impact of flood disasters on child education in Muzarabani District, Zimbabwe. *Jàmbá: Journal of Disaster Risk Studies*, *6*(1), 8.

Munro, A., Kovats, R. S., Rubin, G. J., Waite, T. D., Bone, A., Armstrong, B., Beck, C. R., Amlôt, R., Leonardi, G., & Oliver, I. (2017). Effect of evacuation and displacement on the association between flooding and mental health outcomes: a cross-sectional analysis of UK survey data. *The Lancet Planetary Health*, *1*(4), e134–e141.

Nguyen, C. V., & Minh Pham, N. (2018). The impact of natural disasters on children's education: Comparative evidence from Ethiopia, India, Peru, and Vietnam. *Review of Development Economics*, *22*(4), 1561–1589.

Onigbinde-Jacob, L. (2018). *The impacts of natural disasters on educational attainment: Cross-country evidence from macro data*. Social Science Research Network.

Pelli, M., & Tschopp, J. (2025). Storms, early education and human capital. *Journal of Environmental Economics and Management*, *130*, 103104.

Rannie, W. (2016). The 1997 flood event in the Red River basin: Causes, assessment and damages. *Canadian Water Resources Journal/Revue Canadienne Des Ressources Hydriques*, *41*(1–2), 45–55.

Rush, J. V. (2018). The impact of natural disasters on education in Indonesia. *Economics of Disasters and Climate Change*, *2*, 137–158.

Saif-Ur-Rehman, S. A., & Shaukat, B. (2013). The effects of 2010 flood on educational institutions and children schooling in Khyber Pukhtoonkhwa: a study of Charsadda and Swat Districts. *International Journal of Environment, Ecology, Family and Urban Studies*, *3*(3), 1–12.

Segarra-Alméstica, E., Caraballo-Cueto, J., Cordero, Y., & Cordero, H. (2022). The effect of consecutive disasters on educational outcomes. *International Journal of Disaster Risk Reduction*, *83*, 103398.

Shidiqi, K. A., Di Paolo, A., and Choi, A. (2023). Earthquake exposure and schooling: Impacts and mechanisms. Economics of Education Review, 94:102397.

Shokri, A., Sabzevari, S., & Hashemi, S. A. (2020). Impacts of flood on health of Iranian population: Infectious diseases with an emphasis on parasitic infections. *Parasite Epidemiology and Control*, *9*, e00144.

Thamtanajit, K. (2020). The impacts of natural disaster on student achievement: Evidence from severe floods in Thailand. *The Journal of Developing Areas*, *54*(4).

Tong, S. (2017). Flooding-related displacement and mental health. *The Lancet Planetary Health*, *1*(4), e124–e125.

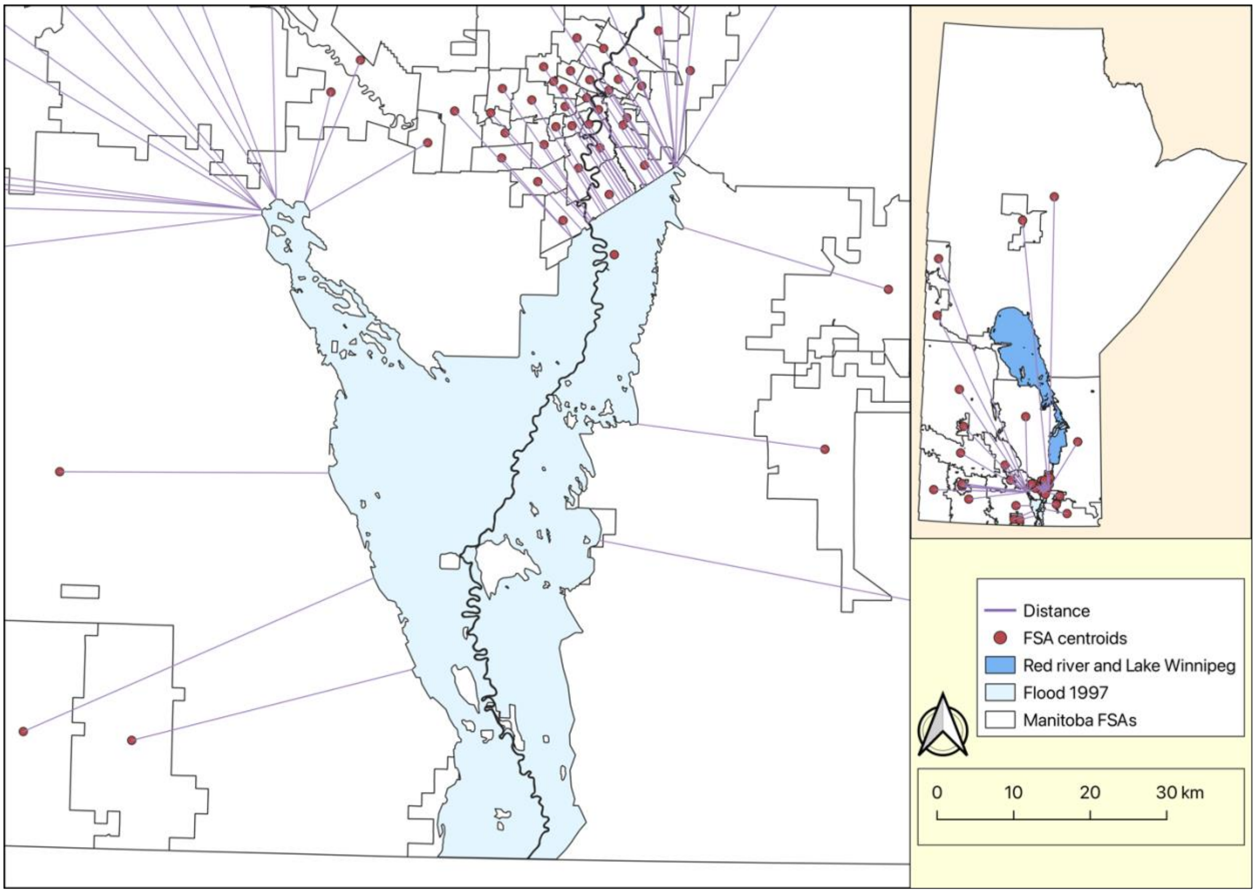
Wang, J. (2024). Impact of natural disasters on student enrollment in higher education programs: A systematic review. *Heliyon*, *10*(6), e27705.

Whitfield, P. H. (2012). Floods in future climates: a review. *Journal of Flood Risk Management*, *5*(4), 336-365.

Figures and Tables

Figure 8: 1997 Red river flood plain

Panel (a): Distance between overland flood boundary and FSA centroids



Panel (b): Treatment and control group classification

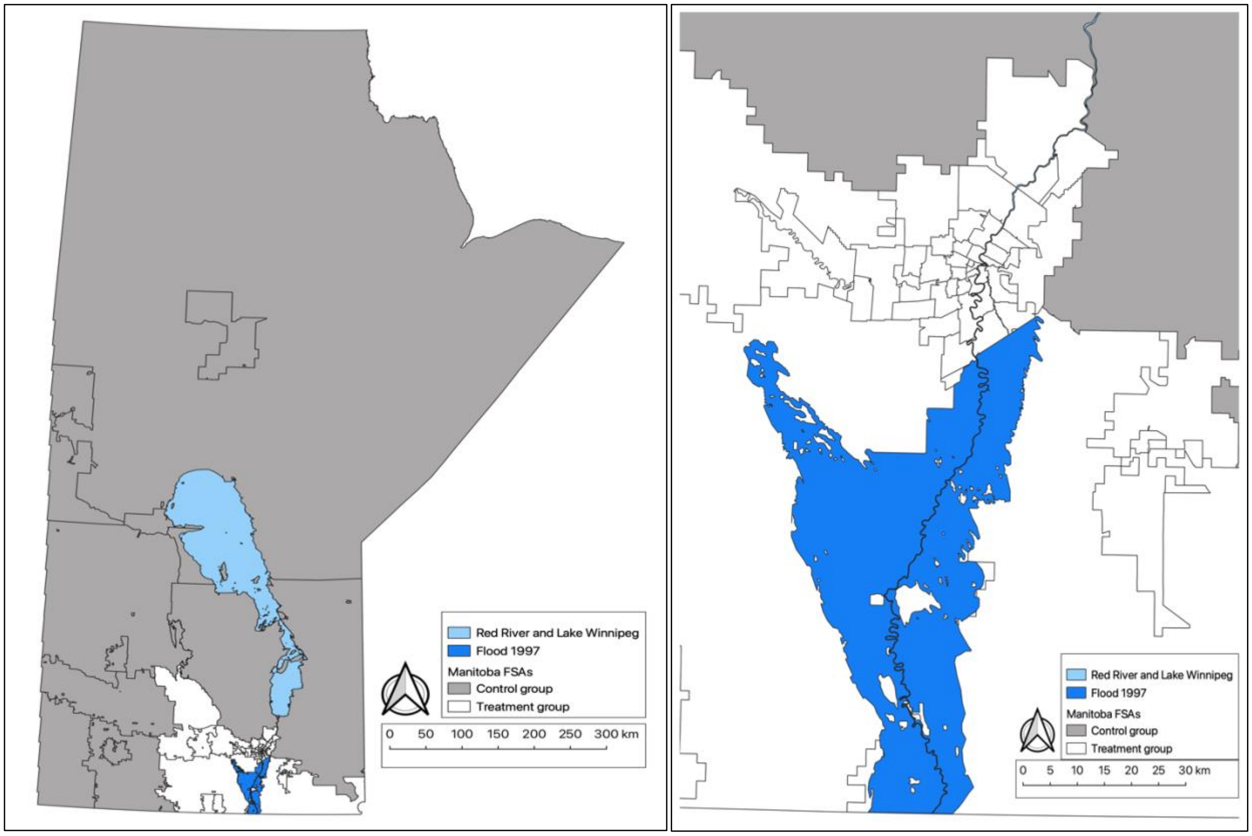


Table 13: Summary statistics

Variable	Mean	SD	Frequency	Percent
Distance	136.428	193.179	14,582	-
Age	6.190	4.861	14,582	-
Motor and social development score	99.856	14.727	5,133	-
PPVT score	100.913	37.760	2,560	-
Math score	431.038	114.935	3,623	-
Female	-	-	7,056	48.39
Male	-	-	7,526	51.61
White	-	-	9,439	64.73
Non-white	-	-	5,143	35.27
Household income < 40 thousand	-	-	4,876	33.44
Household income > 40 thousand	-	-	9,706	66.56

Table 14: Effect of flood on test scores

	(1) Average test score	(2) Female	(3) Male
Flood × Distance	0.385***	0.422**	0.342**
	(0.125)	(0.188)	(0.165)
Observations	11,310	5,572	5,738
Individual control	Y	Y	Y
Household control	Y	Y	Y
Community FE	Y	Y	Y
Year FE	Y	Y	Y

Notes: The dependent variable in Table 14 is the average of three age-specific test scores for each child in each cycle. All estimates are based on equation (1), which includes demographic control variables age, sex, and health condition of the child, parents' education level, household income, community fixed effect (forward sortation area of the child's residence), and year fixed effect (the year in which the survey was conducted). Standard errors in parentheses are clustered at the community level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Effect of flood on test scores by age

	(1) Age < 4	(2) 10 > Age > = 4	(3) Age > = 10
Flood × Distance	0.074**	1.091***	1.283***
	(0.030)	(0.341)	(0.389)
Observations	5,127	4,050	2,133

Notes: The dependent variable in Table 15 is the average of three age-specific test scores for each child in each cycle. All estimates are based on equation (1), which includes demographic control variables age, sex, and health condition of the child, parents' education level, household income, community fixed effect (forward sortation area of the child's residence), and year fixed effect (the year in which the survey was conducted). Standard errors in parentheses are clustered at the community level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Effect of flood on test scores by race

	(1) White	(2) Non-white
Flood × Distance	0.166	0.933***
	(0.102)	(0.257)
Observations	7,522	3,788

Notes: The dependent variable in Table 16 is the average of three age-specific test scores for each child in each cycle. All estimates are based on equation (1), which includes demographic control variables age, sex, and health condition of the child, parents' education level, household income, community fixed effect (forward sortation area of the child's residence), and year fixed effect (the year in which the survey was conducted). Standard errors in parentheses are clustered at the community level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Effect of flood on test scores- learning disability

	(1) Disable	(2) Not-disable	(3) Interaction
Flood × Distance	5.357	0.384^{***}	-
	(6.499)	(0.128)	
Disability x Distance	-	-	0.377^{***}
			(0.133)
Observations	82	11,228	11,310

Notes: The dependent variable in Table 17 is the average of three age-specific test scores for each child in each cycle. All estimates are based on equation (1), which includes demographic control variables age, sex, and health condition of the child, parents' education level, household income, community fixed effect (forward sortation area of the child's residence), and year fixed effect (the year in which the survey was conducted). Standard errors in parentheses are clustered at the community level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Effect of flood on test scores- alternative radius for treatment variable

	(1) 50 km	(2) 100 km (Preferred)	(3) 250 km	(4) 500 km	(5) 750 km	(6) > 500 km
Flood × Distance	0.247^{**}	0.385^{***}	0.111^{***}	0.091^{***}	-0.008	-0.015
	(0.121)	(0.125)	(0.034)	(0.029)	(0.014)	(0.013)
Observations	11,310	11,310	11,310	11,310	11,310	11,310

Notes: The dependent variable in Table 18 is the average of three age-specific test scores for each child in each cycle. All estimates are based on equation (1), which includes demographic control variables age, sex, and health condition of the child, parents' education level, household income, community fixed effect (forward sortation area of the child's residence), and year fixed effect (the year in which the survey was conducted). Standard errors in parentheses are clustered at the community level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Chapter 3 Temperature and Employee Absenteeism: Evidence from Service Sectors in India

Abstract

We analyze the daily impact of outdoor temperature on employee absence in the education and health sectors using data from the Indian Human Development Survey (IHDS). While we found no effect on medical staff absence rates, our findings indicate that a 1°C increase in temperature leads to a 0.6% rise in teacher absence rates in primary schools. This effect is more pronounced in schools with a higher proportion of female teachers or teachers living far from the school and receiving no financial grants. Access to electricity in schools provides protective effects, but only when power outages are infrequent. The results remain consistent across a wide range of robustness tests.

Keywords: Temperature – Employee absence – climate impacts – climate resilience – mitigation.

1 Introduction

Do hot days affect absenteeism at work? We explore the effect of ambient temperature on service sector absence rates, the variation in effect size across different working environments, and how heat-resilient facilities mitigate these effects in a tropical developing country, India. The country often experiences extreme hot days and heat waves during peak summer months, with temperatures soaring to uncomfortable levels (Rohini et al., 2016; Stern, 2008). With a population of 1.4 billion, human capital is one of India's major economic driving forces. Thus, it is essential to investigate how extreme outdoor temperatures affect day-to-day activities and the productivity of employees in service sectors, who are primarily responsible for contributing to the nation's human capital stock.

The rising temperature pattern threatens the global labor market by reducing the supply of low-skilled workers in South Africa (Shayegh et al., 2021) and decreasing the availability of high-risk workers in the Middle East (Abou-Ali et al., 2023). Additionally, there is a projected decline in labor productivity of about 1.6% by 2080 in Europe (Szewczyk et al., 2021) and a reduction in total labor hours from 25.9% to 18.6% in Asia due to a global warming scenario in the near future (Dasgupta et al., 2021). Previous research highlights the significant impact of temperature on economic activities in different countries, such as worker productivity (Zivin & Neidell, 2012), plant-level productivity (Zhang et al., 2018), decision-making (Heyes & Saberian, 2019), annual income growth (Deryugina & Hsiang, 2014), sector-specific output growth, and political stability (Dell et al., 2012). For instance, Park (2017) finds that the negative effects of temperature on labor-intensive products in the United States are more pronounced in industries exposed to extreme heat; however, counties with extreme heat exposure experience less impact than those in cooler regions. The first published study examining temperature sensitivity in the workplace in India provides evidence of labor absenteeism on hot days (Somanathan et al., 2021). Using survey data from selected garment and steel firms, the authors report heterogeneous effects of temperature even in climate-controlled workplaces. They find that absenteeism rises by 0.1 across 224 climate-controlled sewing lines when the same-day temperature exceeds 35°C. Furthermore, when the number of days surpassing this threshold increases in the preceding week, absenteeism rises by 0.045 among nine steel-worker teams and by 0.005 among 147 weaving workers. However, they do not find any immediate or lagged effects of temperature on non-climate-controlled garment facilities.

Leonard (2018), who motivates our study, finds an inverted U-shaped relationship between temperature and teacher presence in India using data from 4,085 schools. Considering 20°C to 25°C as the comfortable temperature range, the author shows that a 1°C temperature increase above the comfortable range raises the teacher absence rate by 8.1%. The most recent study Heyes & Saberian (2022) broadly discusses the link between temperature and the ability to work in the context of India. Based on a nationally representative sample of households, the authors provide evidence that a hot day when the temperature exceeds 37.7°C (100°F) increases the self-reported inability to work on average by 7%. The effect persists for working age groups between 25 and 45 years and above 45 years, as well as among agricultural and construction workers. However, the authors do not find any such effect among respondents employed in official settings.

Another contemporary study by Gupta & Somanathan (2023) investigates the impact of hot days on workplace absenteeism among 200 salaried employees at a large welding company across 86 locations in India. They define absenteeism as the number of days a worker is absent per 1,000 days and conclude that the probability of missing work increases by 4.7% for each degree rise in daily maximum temperature without climate control technology. The lagged effect is more than double, considering the number of hot days in the preceding six days and the specific day of absence, rather than just the daily temperature.

Given the diverse impacts of temperature on worker productivity across different occupational settings, this study explores a similar research question within the same area but in a different professional context involving school teachers and medical facility staff, who generally have greater job responsibility and accountability. Specifically, this paper examines the effect of ambient temperature on teacher and medical staff absence rates in India.

1.1 Contribution

This chapter contributes to the broader literature examining how temperature affects labor supply and productivity, particularly in the service sector. A substantial body of research documents that high temperatures can reduce labor supply, lower productivity, and influence economic output, with especially strong effects in outdoor or physically demanding occupations (Graff Zivin & Neidell, 2014; Hsiang, 2010; Cachon et al., 2012). More recent work has expanded this evidence to indoor, service-oriented roles, showing that even non-manual workers are not immune to heat-related performance declines (Somanathan et al., 2021; Kjellstrom et al., 2016). However, the majority of studies focus on aggregate economic indicators or specific industrial sectors, leaving public education and healthcare underexplored—despite their centrality to human capital formation and societal well-being.

Our study addresses this gap by analyzing the impact of daily maximum temperature on absenteeism in two critical service sectors, schools and healthcare facilities, using nationally representative Indian Human Development Survey (IHDS-II) data. Unlike much of the literature, which often relies on annual or monthly temperature averages, we use fine-grained, daily variation in observed maximum temperature on survey interview dates. This higher frequency allows us to identify short-term behavioral responses to temperature changes that are masked in aggregated data.

In the education context, previous work such as Leonard (2018) has analyzed IHDS-II data at the individual level, using information from 4,085 schools across 355 districts in India to determine whether teacher presence decreases on hotter days when temperatures exceed 25°C. Leonard considered teacher presence as a binary outcome, and modeled this relationship using OLS, logistic, and probit regressions, finding a non-linear inverted U-shaped relationship between temperature and teacher presence, with all temperature bins above 25°C–30°C being significant. However, methodological limitations (including coding errors and selective sampling of hot days) raise concerns about bias and generalizability. We improve upon this line of research by:

1. Incorporating the full temperature distribution rather than imposing an arbitrary threshold of >25°C, which avoids selection bias, allows for proper randomization, and better captures heterogeneous heat sensitivity among workers.
2. Aggregating absence data to the school level to reflect systemic service delivery impacts, capture the overall educational impact on students, and allow for stronger policy relevance rather than treating attendance as an individual binary outcome.
3. Incorporating a larger dataset, with information from 4,100 schools across 371 districts, which makes the estimates more precise and representative of the national context.
4. Comparing the effect of temperature on two service sectors—teachers and healthcare providers—which reveals that while higher temperatures significantly increase teacher absence rates, healthcare absence rates are unaffected, suggesting that better facilities and stronger professional obligations may mitigate heat effects.

Our findings differ from Leonard's in that we do not detect a significant teacher-level effect in the 20°C–25°C range and do not confirm the inverted U-shape; rather, we find a significant school-level effect for daily maximum temperatures exceeding 30°C–35°C and a linear relationship beyond 20°C–25°C. Moreover, by contrasting education with healthcare, we reveal that temperature has no statistically significant effect on healthcare worker absenteeism, suggesting that environmental controls, job design, and institutional norms in healthcare may offer protection against climate-related productivity losses.

In doing so, this chapter not only corrects and refines prior education-sector evidence but also contributes to the general labor supply literature by showing that the magnitude and even the direction of temperature effects are sector-specific. This has important implications for climate adaptation policies, highlighting that investments in workplace environment and organizational resilience can substantially mitigate the adverse labor supply impacts of extreme heat, even in resource-constrained settings.

2 Education System in India

Primary or elementary education in India encompasses the schooling of children aged 6 to 14 years (Bajpai & Goyal, 2004). In 2009-2010, 99% of the rural population in India had a primary school located within 1 km of their residence. However, only 53% of schools maintained the recommended 1:30 student-teacher ratio at the primary level. In contrast, others had ratios as high as 46:1. Recruitment procedures and training terms of teachers vary from state to state (Bordoloi, 2011). Hiring through social connections, nepotism, and bribery is common in the Indian school system (Bajpai & Goyal, 2004). Deployment and transfers are subject to significant bureaucratic and political intervention. In 2002-2003, India had over 15% of schools operated by a single teacher, with a four times higher percentage in rural areas than urban areas (Govinda & Bandyopadhyay, 2008). Poor basic facilities in schools, such as one or two classrooms, multi-grade classrooms, and a lack of running water or toilets (Bajpai & Goyal, 2004), along with poor residential facilities in rural areas, contribute to lower pupil-teacher ratios (Govinda & Bandyopadhyay, 2008). India has the second highest teacher absence rate among eight countries that use a similar methodology for absence calculation. One in four teachers in government primary schools is absent on a typical day (Kremer et al., 2005). Teachers frequently remain absent on school days without prior notice due to inadequate and irregular supervision (Bordoloi, 2011). Higher salaries for teachers do not effectively incentivize better attendance or class performance (Kremer et al., 2005; Grover & Singh, 2002). Moreover, teachers are often not held accountable for their duties due to politically supported teachers' unions, which wield significant influence over changes to the daily school schedule, deviating from the standard administrative opening and closing times. Unofficial school closures due to excessive heat, heavy rainfall, peak harvesting seasons, wedding seasons, and religious festivals are common yearly phenomena (Bajpai & Goyal, 2004).

3 Health Care System in India

India has a diverse healthcare system comprising public and private facilities across primary, secondary, and tertiary levels. The primary level includes sub-centers and primary health centers. Sub-centers are located in remote areas to serve populations of 3,000 to 5,000. Primary health centers serve more developed rural areas with populations of 20,000 to 30,000. The secondary level consists of community health centers designed to serve 80,000 to 120,000 people in urban areas and sub-district hospitals that provide more specialized care. The tertiary level includes district hospitals and medical colleges, which are referral centers for primary and secondary levels (Chokshi et al., 2016).

The country faces a substantial shortage of healthcare professionals across all three levels and offers only one-fourth of the World Health Organization's recommended 2.3 medical staff per 1,000 people (Rao et al., 2012). Urban areas are more densely staffed with healthcare workers than rural areas (Singh et al., 2014; Saxena et al., 2015). While citizens receive free health care services at public facilities, a shortage of staff and medical supplies, especially in rural areas, remains a major concern.[24] This forces people to seek medical care at private facilities at a higher cost, which increases their financial burden because over 80% of the population does not have health insurance coverage in India (Rudrappa et al., 2018). Additionally, employees at public health facilities typically experience less job stress and show less commitment to their organization compared to employees at private health facilities (Kumar et al., 2013; Kumar et al., 2014). They take more paid time off without official reasons, which increases government spending on lower-quality services. Staff shortages and inefficient service have led to the rapid growth of private healthcare facilities in India (Hammer et al., 2007). Consequently, private health services dominate India's dynamic medical sector, offering major secondary and tertiary-level facilities mainly concentrated in larger cities. Currently, more than 57% of citizens in India prefer private institutions for medical care over government-run facilities. However, despite numerous drawbacks, India's healthcare system remains the largest revenue-generating sector, leveraging its competitive advantage of a

[24] Detail at https://web.archive.org/web/20200214134709/http://nhsrcindia.org/sites/default/files/NHA%20Estimates%20Report%20-%20November%202018.pdf

large pool of skilled medical professionals and low-cost services compared to other Asian countries (Rudrappa et al., 2018).

4 Data

4.1 India Human Development Survey

The India Human Development Survey (IHDS-II, 2011–2012) is a nationally representative survey that includes 31,994 teachers, 4,267 primary schools across 384 districts, 1,503 villages, and 971 urban neighborhoods in India. The survey covers diverse household perspectives, including socio-demographic status, economic conditions, employment records, health, and school facilities. To compare temperature effects across service sectors, we focus particularly on the education and health modules of IHDS-II. The health module covers 4,447 medical facilities and 23,327 medical staff within the same geographic area.

This survey was conducted by the University of Maryland and funded by the National Institutes of Health (NIH), the United States Department of Health and Human Services (HHS), and the Ford Foundation. The IHDS maintains high-quality data according to established metrics (Wang et al., 2014). A highly experienced team of interviewers from the National Council of Applied Economic Research (NCAER) carried out pre-scheduled interviews in a randomized order from January 2011 to December 2012. The NCAER team interviewed up to two randomly chosen primary schools and medical facilities (one government and one private) located within the sample villages or urban blocks selected through stratified random sampling. If a school or a medical facility was not found in the selected village, the next nearest school or medical facility was surveyed.

The school and medical facility datasets offer details on the school and clinic environments, teaching performance, and health care services. The staff datasets include a comprehensive set of questions about staff members' age, sex, education, religion, reasons for absence, distance between workplace and home, and more. At the school level, the survey provides information on the type of school, grant status, primary water source, connection to an electricity grid, quality of power supply, ownership of cooling technology, and washroom facilities. Facility-level data includes information on facility type, services provided, availability of necessary medical equipment, daily hours of operation, number of patients served, and other relevant details.

Our outcome variable is based on the question, *'Is Name X present today at the school or the clinic?'* First, we create a count variable for the total number of teachers or medical staff absent out of all teachers and medical staff interviewed at school or medical facility i on the interview date t. Then, we calculate the absence rate for each school and each medical facility on that specific interview date. We exclude all absences due to government work, official obligations, census activities, or election duties. After dropping duplicates and merging with weather data, we have 4,100 schools and 2,747 medical facilities for analysis.

Our data shows that 2,201 of 4,100 schools have more than 60% female teachers, while 1,903 of 2,747 medical facilities have more than 60% male staff. Figure 9, Panels (a) and (b) show the district-wise average absence rate for primary schools and medical facilities. The mean of district-wise average teacher absence rate is 12.5%, with a standard deviation of 6.3%. *Jabalpur* district in *Madhya Pradesh* has the highest teacher absence rate at 82.6%. Most of these schools have 100% female teachers and experience maximum daily temperatures exceeding 40 degrees Celsius. In contrast, schools in *Deoria* and *Baliya, Uttar Pradesh,* have zero absence rates, enjoy pleasant weather with daily highs ranging from 20 to 22 degrees Celsius, and nearly all—except for one—have more male teachers. On the other hand, the mean of district-wise average medical staff absence rate is 6.2%, with a standard deviation of 2.7%. *Deoria* district in *Uttar Pradesh* has the highest medical staff absence rate at 30%, with 80% male staff, and the maximum daily temperature between 20 and 23 degrees Celsius.

Figure 11, Panel (a) displays a basic scatter plot of district-level teacher absence rates and temperature, suggesting a linear relationship between the two variables. In contrast, Panel (b) shows a more dispersed pattern between district-level medical staff absence rates and temperature.

4.2 Weather

We face two distinct challenges when using climate data from ground-based public stations in India: 1) the exact location of each school or medical facility is unknown, making it difficult to assign weather data from the nearest station, and 2) the spatial and temporal coverage of these stations is inadequate. Therefore, following Schlenker & Lobell (2010), Schlenker & Roberts (2009), and Auffhammer et al. (2013), we utilize the reanalysis weather datasets from the ERA-interim archive of the European Centre for Medium-Range Weather Forecasts (ECMWF).[25] Specifically, we obtain daily maximum and minimum temperatures, precipitation, relative humidity, and solar radiation data from January 1, 2011, to December 31, 2012. The satellite raster datasets are presented as $1^{\circ} \times 1^{\circ}$ latitude-longitude grid cells. The geographic granularity of the IHDS-II data is limited to the district level.[26] Therefore, we compute the inverse distance-weighted average of each weather variable across all grid points falling within 20 miles of the centroid of

[25] Details about ERA-interim are available at https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets
[26] A district is the third tier in a six-tier administrative system in India and is officially known as revenue districts. The country is divided into 35 states and 593 districts (Census of India, 2001).
Details at https://censusindia.gov.in/census.website/data/atlas

each district polygon. The final dataset provides a panel of weather variables at the district-date level.

Figure 10 displays the temperature distribution during *summer*, *fall*, and *winter*. In our sample, the average of daily maximum temperature is 31° Celsius with a standard deviation of 7°C. In *summer,* the average is 33°C with a standard deviation of 8.5°C. In April and May, the median temperature is 37°C; in June and July, it is 31.5°C and 28.5°C, respectively.

5 Method

We utilize both a continuous and a non-linear measure of daily maximum temperature as the treatment of interest to examine its concurrent effect on unplanned teacher or medical staff absenteeism. We begin with the following fixed-effect specification:

$$A_{idt} = \alpha + temp_{idt}\beta + W_{idt}\theta + S_{idt}\gamma + FE_{DoW} + FE_{state} + FE_{ymd} + \epsilon_{idt} \qquad (1)$$

where A_{idt} is the percentage of teachers or medical staff absent in school or medical facility i in district d on the interview date t. The treatment variable $temp_{idt}$ is the observed value of the maximum daily temperature in school or medical facility i, which is identical to that of the other schools or medical facilities in the respective district d.

The majority of the states in India experience a mixed tropical climate (i.e., dry, wet, and humid) along with regional variations in rainfall which might have an effect on institutional activities. Therefore, we include a vector of weather controls W_{idt}, considering the confounding role of precipitation, humidity, and solar radiation in our setting. The vector $S_{idt}\gamma$ includes school controls such as private-public status, school age, student enrollment, student attendance on an average day, number of classrooms, number of full-time teachers, number of part-time teachers, number of shifts in school. Facility-level controls include private-public status, facility age, facility type, average beds occupied at night, average outpatients each week, number of staff working at the clinic, number of vacant sanctioned positions, and total hours of operation per week.

Variation in the response may also stem from several spatial and temporal differences, such as the timing and pace of adopting a new education or health policy, unexpected natural or political disasters, cultural holidays, or differences in the academic calendar.[27] We include different combinations of several fixed effects to account for these unobservable phenomena. Specifically, our preferred specifications incorporate state (FE_{state}), day-of-week (FE_{DoW}), and year-month-district (FE_{ymd}) fixed effects.

ϵ_{idt} is the cluster-adjusted standard error. We cluster standard errors at the district level in all specifications to capture within-district correlation for temperature assignment, following the methodologies outlined by Heyes & Saberian (2019) and Abadie et al. (2023).

5.1 Identification

Our identification strategy exploits short-term variation in maximum daily temperature across districts to estimate its effect on unplanned teacher or medical staff absenteeism. These temperature fluctuations are driven by naturally occurring meteorological processes such as air mass movements, frontal systems, and seasonal shifts that are determined by large-scale atmospheric dynamics and cannot be influenced by local socio-economic conditions or institutional behaviour in the short run. This makes them a plausibly exogenous source of variation for causal inference once we control for location-specific climate norms and common time shocks.

[27] Summer break in India typically lasts for 6 to 8 weeks. Schools in the northern states remain closed from mid to late May until early July, while schools in southern states close in mid-April and reopen in late May or early June.

We measure temperature using the inverse-distance weighted average of maximum daily readings from all grid points within 20 miles of each district centroid, assigning the same value to all institutions in the same district on a given day. Because the locations of geostationary satellites are fixed and their placement is unrelated to institutional or community-level absenteeism patterns, the assignment of temperature is orthogonal to unobserved determinants of the outcomes. In our preferred specification, variation in assigned temperature within a district over time is assumed to be as good as random after accounting for other weather variables.

The main threats to identification are confounding weather factors and spatial or temporal shocks correlated with both temperature and absenteeism. To address these, we control for precipitation, humidity, and solar radiation, which may directly influence both institutional functioning and absenteeism. We also incorporate school- and facility-level characteristics, such as ownership, size, staffing levels, and operating conditions, to account for differences in institutional capacity and baseline attendance patterns.

Another challenge for our estimation is residential sorting or location choice endogeneity. For instance, individuals sensitive to humid summers might choose to live in areas with a temperate climate (Heyes & Saberian, 2019; Greenstone & Chay, 2005; Graff Zivin & Neidell, 2009). However, the temperature assignment in our method is conditional on the district level, which assumes relocation happens mostly within a district. Thus, the variation in assigned temperature after all locational adjustments is expected to be sufficiently random to allow for true causal inference in our preferred specification

We further mitigate omitted-variable bias by including a rich set of fixed effects:

- **State fixed effects** control for time-invariant state-level differences, such as governance structures, long-run climate, and infrastructure quality.

- **Day-of-week fixed effects** account for systematic weekly absenteeism patterns.
- **Year–month–district fixed effects** absorb location-specific seasonal patterns and time-varying district-level shocks (e.g., academic calendar changes, local festivals, regional policy shifts).

These fixed effects remove the influence of unobserved, time-invariant spatial heterogeneity and slow-moving or recurring temporal shocks, isolating plausibly exogenous temperature-driven variation.

Our identification assumptions are likely to hold for three reasons. First, day-to-day and year-to-year temperature variation within a district is unpredictable and uncorrelated with unobserved determinants of absenteeism once weather controls and fixed effects are included. Second, the fixed location of the geostationary satellites and district-level temperature assignment rules out measurement endogeneity from relocation, especially since most residential mobility occurs within districts. Third, prior research Heyes & Saberian (2019); Greenstone & Chay (2005); Graff Zivin & Neidell (2009) has validated the use of short-run weather variation as a quasi-experimental source of exogenous shocks in similar labour and health settings.

6 Result

6.1 Main Result

Table 20 presents the effects of daily temperature variation on teacher absences at the school level in India under different specifications. Starting at column (1), the estimate without controls indicates that the absence rate is positive and significant at 1%. Columns (2) and (3) report slightly smaller estimates with weather and school controls, but retain the same significance level. In columns (4) through (8), we incorporate day-of-week, district, year, year-month, and state-fixed effects. Column (9) shows our preferred specification, which includes day-of-week, state, and year-month-district fixed effects while excluding their individual spatiotemporal effects. This approach helps us mitigate any potential concern of multicollinearity arising from including similar variables. Our inference from column (9) is that a 1° Celsius rise in temperature increases the teacher absence rate by an average of 0.6 percent, holding all else constant, and is significant at 5%. Although our preferred result is less significant than earlier specifications, it allows us to achieve relatively stable estimates while accounting for all location and time fixed effects. Table 21 estimates the same models with the medical staff dataset. None of the estimates shows a significant effect of temperature on the medical staff absence rate. This is expected as health care providers play a critical role in people's health and are less likely to be absent from work due to weather changes. They also work in a better environment that is conducive to their higher job responsibilities and work ethic.

Our goal is to quantify the immediate impact of daily temperature variation, especially at a time when daily weather forecasts are accessible to everyone. Table 22 compares our main findings with the lag and lead specifications. We consecutively incorporate two lags and two leads of the daily maximum temperature. In columns (2) to (5), we do not observe any effect of the previous two days' temperatures on teacher absence, confirming the validity of our main specification. However, our estimates reveal significant negative effects of adding the temperature of the next two days following the interview date, which is expected. Table 23 estimates the exact specifications with temperature lags and leads for the medical staff dataset. We find some significant positive estimates with 2-day temperature lags and leads, yet we do not conclude anything from this result because of the lower sample sizes in these models.

6.2 Age and Sex

Previous research on temperature indicates that older adults and females are more susceptible to heat exposure than younger adults and males based on various epidemiological measures, including mortality (Deschenes, 2018; Deschênes & Greenstone, 2011), morbidity (Barreca & Shimshack, 2012; Liss et al., 2017), and mental health (Heyes & Saberian, 2019; Dai et al., 2016).

We re-estimate the preferred model across different sub-samples to reassess whether individual traits influence the effect of heat exposure. We categorize schools based on their sex composition rate and the age threshold at the 25th percentile (32 years). Generally, most schools in our sample have a higher female participation rate and more teachers over 32 years old. Columns (1) and (2) in Table 24 report estimates for schools with 50% or more female and male participation, respectively. We find that the teacher absence rate significantly increases by 0.6% in schools with more female teachers, older female teachers, and a greater number of older teachers overall. In contrast, schools with more male teachers and older male teachers remain unaffected by the rise in ambient temperature.

Our results show that heat exposure affects work performance for female teachers more than male teachers. These findings align with physiological literature that suggests men are more heat resistant than women due to physical and biological differences, such as cardiovascular fitness, body size, heat acclimation, heat storage, and sweating sensitivity (Kenney, 1985; Burse, 1979; Bittel & Henane, 1975). Dry or humid weather disrupts women's daily activities in various ways, including migraines (Yang et al., 2015; Scheidt et al., 2013), fainting (Chaudhary et al., 2024; Grubenhoff et al., 2007), and physical exhaustion (Glazer, 2005). Moreover, heat stress, especially when combined with humidity, reduces evaporative capacity (Lei et al., 2017).

Therefore, our findings are particularly relevant for India, where women experience both indoor (cooking and household chores) and outdoor heat stress while often balancing dual roles at home and in the workplace.

We do not have the medical staff's age information in the dataset. Table 25 shows estimates based on the male and female staff rates at medical facilities. However, we have not found any significant result that validates the high work ethic of employees in the medical sector, irrespective of sex.

6.3 Distance from school

Distance from school is crucial in a developing country like India.[28] In our sample, 50% of the teachers live within 1 kilometer, and 75% live within 5 kilometers of the school premises, with an average distance of 4.2 kilometers. Therefore, to explore whether distance affects the absence rate alongside temperature, we define teachers living more than a kilometer away from school as '*away*.' Next, we generate the percentage of teachers in each school who live away.

Column (1) in Table 26 illustrates that temperature and living away from school significantly affect the absence rate. Focusing only on extremely hot days when the daily temperature exceeds 29 °C, a 1% increase in the number of teachers living away raises the teacher absence rate by an average of 0.04%. We introduce additional specifications in columns (3) to (6) to ensure that schools in our sample with all teachers a*way* do not influence our estimates. Columns (3) and (5) represent results from columns (1) and (2), excluding the top 10% of schools where all

[28] Around 60% of kids, with a higher proportion for primary and upper-primary children, walk on foot every day to attend school (The Times of India, 2020). We do not have statistics for primary school teachers, but our assumption is that individuals, by nature, would prefer to live closer to the workplace (for instance, within a km, which is 15-20 minutes walking distance) and commute on foot in a developing country's setting. Our definition of being 'away' considers 1 km cut off as half of the teachers in the sample live within a kilometer of school premises.

teachers are *away*. Columns (4) and (6) show corresponding estimates, excluding the top 25% of schools where 75% of the teachers live far away. In conclusion, we find a significant impact of temperature and distance from school, even after excluding teachers who live far away.

We perform the same exercise in Table 27 using the medical staff dataset. As expected, we do not observe any effect of temperature, but we have found significant effects of 'living away' on the staff absence rate.

6.4 Seasonal variation

As a South Asian country, India experiences changing weather patterns across seasons. Therefore, we expect the impact of temperature to vary across India's four official seasons. Figure 12 displays estimates for the entire sample during *'winter,' 'summer,' 'monsoon*,' and *'post-monsoon*.' The effects of temperature on the absence rate show an upward trend as *summer* and *post-monsoon* progress. However, the effect is significant only during the *summer* season. These findings are reasonable, as *winter* is typically cold, the *monsoon* brings substantial rainfall to South Asia, cooling the weather, and the *post-monsoon* is not as hot and humid as *summer*. Although the estimate for *post-monsoon* is insignificant, the upward trend may be linked to major festivals occurring in October and November, when most schools remain closed.

7 Mitigation and resilience: Power supply, water, and cooling facilities

7.1 Electricity

The health and economic impacts of climate change, including strokes, suicide, migraine discomfort, reduced work capability, and crop loss, are more severe in developing countries (Hansen et al., 2012). Therefore, addressing climate change on a global scale and mitigating its effects at a national level presents dual policy implications for less developed nations (Barreca et al., 2016). In this section, we examine whether school facilities—such as access to electricity, the quality of the power supply, and the installation of fans—can mitigate the estimated effects of hot weather in our context.

Table 28 demonstrates a subsample analysis considering access to electricity. Temperature estimates in columns (1) and (2) are significant and almost equal in magnitude. However, both the electricity control and the interaction between electricity and temperature are insignificant. Therefore, being connected to a power grid does not affect the teacher absence rate. Columns (3) and (4) show that the absence rate increases with temperature in schools that have electricity for extended periods. It is reasonable to assume that the length of the power supply is non-random and depends on the neighborhood wealth measure of a school. Teachers from less affluent regions may be better adapted to hot weather without a sufficient power supply, as column (3) indicates, compared to teachers from affluent communities. We cannot compare connected and unconnected schools since most of the schools in our sample have electricity connections.

7.2 Quality of power supply

Power failures cause economic losses in low-income countries, disrupting the irrigation, business, health, and education sectors. Although India is one of the largest providers of energy subsidies, maintaining a continuous power supply remains a significant challenge, particularly during the hot summer months.[29] Here, we explore the importance of a stable power supply and its connection to the power grid.

[29] Details in UNEP (2008), available at https://www.cbd.int/doc/case-studies/inc/cs-inc-unep-reforming-energy-subsidies.pdf

In IHDS-II, 49% of schools experience daily power outages. Column (2) of Table 29, including the outage variable as a control, increases the estimated coefficient by 0.01%. We do not observe any distinct or interactive effects between power outages and temperature. However, the absence rate rises by an average of 0.7% in schools that experience daily power outages. Conversely, the effect is not significant for schools that face power cuts infrequently or not at all.

7.3 Fan ownership

Cooling equipment plays a crucial role in mitigating the impacts of climate change, such as preventing heat stress and illness, improving learning outcomes, and enhancing exam performance (O'Neill et al., 2009; Goodman et al., 2018; and Park, 2017). We assume that the only cooling devices available in schools are electric fans since IHDS-II includes explicit questions regarding the presence of fans in the principal's office and classrooms. About 32% of the schools in our sample do not have a fan, which we believe are predominantly located in rural areas. We re-estimate the preferred specification by adding an interaction between fan presence and temperature and controlling for fan presence separately in Table 30. None of these are significant and do not affect our temperature estimates. Additionally, we do not observe any effect of temperature in schools, regardless of fan presence.

8 Robustness

In this study, we focus on the daily maximum temperature as the main treatment variable to more accurately portray the actual effect of the realized temperature. Next, we conduct several robustness exercises using alternative measures of temperature to test the validity of our main findings.

8.1 Alternative Measure of Temperature

Previous literature considers average temperature over a specific time window (Heyes & Saberian, 2022) or on a monthly or quarterly basis (Ranson, 2014; Goodman et al., 2018). We follow their approach in a slightly different manner, using the 25th percentile of the daily average temperature (22°C) as the threshold for hot days. The estimated coefficient in column 1 of Table 31 indicates a significant increase in the absence rate by 0.8%. We also estimate the same specification by considering the minimum daily temperature below 13.23°C (25th percentile) in column (2). The coefficient estimate at this low level of temperature is negative, suggesting that temperate weather significantly reduces the absence rate, which supports the seasonal estimates for *winter* and *monsoon* found in our earlier specifications.

Next, we substitute the average temperature with the diurnal temperature range (DTR), which is defined as the difference between the daily maximum and minimum temperatures. DTR has been associated with emergency hospital visits (Davis et al., 2020), mortality (Zhang et al., 2018), and cardiovascular diseases (Zheng et al., 2016) during both cold and hot days. Our DTR analysis for the entire sample shows a 0.3% higher effect than the daily maximum temperature. We then re-estimate the same model, excluding extreme DTR values (> 16 °C) in column (4), leading to an even higher effect of more than 1 %.

Humidity is an essential factor in the perception of heat and should be considered when examining the effect of temperature (Baylis, 2015; Heyes & Saberian, 2019). To test whether humidity influences the effect of DTR, we introduce an interaction between DTR and humidity in column (5). In our full sample analysis, we can directly compare columns (3) and (5). Although we lose significance, the coefficient size of the interaction term remains fairly consistent.

Solar radiation is another vital factor, as DTR increases with rising surface solar radiation (Makowski, 2009). Most days in our sample exhibit low solar radiation values with a mean of 19 W/m^2.[30] Nevertheless, we re-estimate column (3), excluding the top 25% (> 24 W/m^2) of the highest radiation days in our sample, which results in a significant coefficient with a 0.3% higher effect in column (6).

Finally, we include the heat index in column (7), also known as apparent temperature, to compare the effect of daily maximum temperature with perceived temperature. Although half of the interview days in our sample are dry and comfortable (with 55% humidity), we still find evidence of a significant rise in the absence rate when the heat index increases.

Figure 13 displays our preferred specification using a non-linear representation of daily maximum temperature. Here, temperature is treated as a categorical variable divided into eight bins, each five degrees Celsius wide, with the 20°C to 25°C range serving as the reference category. The distribution of coefficients indicates a near-linear trend, which supports the linear specification in our main findings. Given the relatively small amount of data in these bins, coefficients below the reference category have very wide confidence intervals.

9 Additional Findings

In this section, we add further exercises based on our preferred specifications.

In Table 32, we examine whether the location of the class, indoors or outdoors, influences our estimates. We include an interaction between indoor class status and temperature, controlling for indoor class status separately in columns (1) and (2). Estimates on temperature are significant,

[30] Solar radiation is classified as low radiation (less than 2.6 kilowatt-hours per meter squared), moderate radiation (2.6-3 kWh/m^2), high radiation (above 3-4 kWh/m^2), and very high radiation (above 4 kWh/m^2).

though neither the interaction term nor the indoor status affects the absence rate. Most of the schools conduct indoor classes, so the effect of temperature in column (3) remains consistent with the main results.

Table 33 presents the results from sub-sample exercises based on the schools' grant status. Overall, the temperature effect is significant for schools without grants, while this effect is balanced out in schools with grants.

In Table 34, we perform similar analyses using the schools' evaluation criteria. We find that a rise in temperature significantly increases the absence rate in schools with teacher evaluation norms. Thus, schools with evaluation criteria are directly comparable to those without any evaluation criteria. Moreover, most schools evaluate teachers based on criteria other than attendance, which might intensify the effect of temperature.

All of our specifications are clustered at the district level. In Table 35, we substitute the preferred standard errors with alternative standard errors across five different spatiotemporal levels. Columns (2) to (7) display all estimates considering clustered standard errors at the school, school-month, district-year, state-year, year-month, and climate zone-month levels, which are consistent with our preferred estimate. We include the Eicker-Huber-White standard error in column (8), which also aligns closely with the other estimates in this table.

Next, we divide the study area into six distinct climate zones, following Bal et al. (2016). Table 36 shows that the effect of temperature is only significant for the Northwest zone, the hottest region of India, which includes *Delhi, Rajasthan, Punjab, Haryana, Gujarat,* and *Chandigarh*. However, we refrain from drawing any inference based on this finding, as dividing schools into zones results in a much smaller sample size for analysis.

10 Conclusion

In an era of accessible daily weather forecasts, we analyze the causal relationship between temperature and absenteeism in the service sector using primary school and medical staff data from IHDS-II. The coverage of schools and medical facilities provided by IHDS-II is nationally representative, and the outcome variable is the actual daily absence rate for each school or medical facility. This approach ensures responses are free from recall bias, and neither the interview schedule nor the data collection method influences it. We exclude all absences related to census or government obligations to ensure our estimates do not reflect ambiguous effects. Considering the daily maximum temperature as the main treatment variable, we find that a 1° Celsius increase in temperature raises the teacher absence rate by an average of 0.6 percent. Our estimates primarily indicate the immediate effect of temperature rather than any lagged effects, supporting our initial hypothesis about temperature and absenteeism. Our additional exercise reveals that temperature has no significant effect on absence at the teacher level, as opposed to the inverted U-shaped relationship in Leonard (2018). We do not observe any temperature effect on medical staff absence rates. Comparing temperature effects across different service sectors, we provide strong evidence that better working environments, higher job responsibilities, and work ethic mitigate any impact of temperature exposure, highlighting the need for essential measures in sectors with poor working conditions and inadequate regulations.

As extreme temperature patterns become more frequent, life on Earth gradually adapts while facing significant economic and health-related challenges. Therefore, it is essential to understand how and to what extent economic sectors can adjust their normal operations through adaptation and mitigation strategies. We show that fans as a cooling facility do not have any mitigating effect. At the same time, financial grants reduce the impact of temperature, while the quality of the power supply also plays a significant role. Our findings are comparable to those of Somanathan et al. (2021), who suggest minimal to no effect of workplace adaptation on productivity or absenteeism. The observed mitigation effect in our setting has essential implications for understanding the sector-wise relationship between labor supply and climate resilience. Depending on work intensity, different economic sectors may respond differently to various adaptation strategies, which remains an important area for further investigation.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, *138*(1), 1–35.

Abou-Ali, H., Hawash, R., Ali, R., Abdelfattah, Y., & Hassan, M. (2023). Is it getting too hot to work? Investigating the impact of climate change on labour supply in the MENA region. *Applied Economics*, *55*(25), 2847–2862.

Auffhammer, M., Hsiang, S. M., Schlenker, W., & Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, *7*(2), 181-198.

Bal, P. K., Ramachandran, A., Palanivelu, K., Thirumurugan, P., Geetha, R., & Bhaskaran, B. (2016). Climate change projections over India by a downscaling approach using PRECIS. *Asia-Pacific Journal of Atmospheric Sciences*, *52*, 353–369.

Bajpai, N., & Goyal, S. (2004). *Primary education in India: Quality and coverage issues*. (Center on Globalization and Sustainable Development Working Papers No. 11).

Barreca, A., Clay, K., Deschenes, O., Greenstone, M., & Shapiro, J. S. (2016). Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, *124*(1), 105–159.

Barreca, A. I., & Shimshack, J. P. (2012). Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the United States. *American Journal of Epidemiology*, *176*(suppl_7), S114–S122.

Bittel, J., & Henane, Rjtj. (1975). Comparison of thermal exchanges in men and women under neutral and hot conditions. *The Journal of Physiology*, *250*(3), 475–489.

Burse, R. L. (1979). Sex differences in human thermoregulatory response to heat and cold stress. *Human Factors*, *21*(6), 687–699.

Baylis, P. (2015). *Temperature and temperament: Evidence from a billion tweets.* Energy Institute at HAAS Working Paper, 12(1), 1–21.

Bordoloi, R. (2011). Challenges in elementary education in India: various approaches. *Journal of Education and Practice*, *2*(7), 39–45.

Cachon, G. P., Gallino, S., & Olivares, M. (2012). Severe weather and automobile assembly productivity. *Columbia Business School Research Paper*, (12/37).

Chaudhary, P., Bharadwaj, A., Gupta, R. K., Kumar, S., & Sharma, S. (2024). Heat Stress and Dehydration. In Adaptation under Stressful Environments through Biological Adjustments and Interventions. *Springer*, pp. 369–376.

Chokshi, M., Patil, B., Khanna, R., Neogi, S. B., Sharma, J., Paul, V. K., & Zodpey, S. (2016). Health systems in India. *Journal of Perinatology*, *36*(3), S9-S12.

Dai, L., Kloog, I., Coull, B. A., Sparrow, D., Spiro III, A., Vokonas, P. S., & Schwartz, J. D. (2016). Cognitive function and short-term exposure to residential air temperature: A repeated measures study based on spatiotemporal estimates of temperature. *Environmental Research*, *150*, 446–451.

Dasgupta, S., van Maanen, N., Gosling, S. N., Piontek, F., Otto, C., & Schleussner, C.-F. (2021). Effects of climate change on combined labour productivity and supply: an empirical, multi-model study. *The Lancet Planetary Health*, *5*(7), e455–e465.

Davis, R. E., Hondula, D. M., & Sharif, H. (2020). Examining the diurnal temperature range enigma: why is human health related to the daily change in temperature? *International Journal of Biometeorology*, *64*(3), 397–407.

Dell, M., Jones, B. F., & Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, *4*(3), 66–95.

Deryugina, T., & Hsiang, S. M. (2014). *Does the environment still matter? Daily temperature and income in the United States*. NBER Working Papers 20750, National Bureau of Economic Research, Inc.

Deschenes, O. (2018). Temperature variability and mortality: Evidence from 16 Asian countries. *Asian Development Review*, *35*(2), 1–30.

Deschênes, O., & Greenstone, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, *3*(4), 152–185.

Glazer, J. L. (2005). Management of heatstroke and heat exhaustion. *American Family Physician*, *71*(11), 2133–2140.

Goodman, J., Hurwitz, M., Park, J., & Smith, J. (2018). Heat and learning. *American Economic Journal: Economic Policy* 12 (2): 306–39.

Govinda, R., & Bandyopadhyay, M. (2008). Access to Elementary Education in India. *Country Analytical Review.* ERIC.

Graff Zivin, J., & Neidell, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, *32*(1), 1-26.

Graff Zivin, J., & Neidell, M. (2009). Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management*, *58*(2), 119–128.

Grammarly. (2025). Grammarly (Version 14.1232.0) [Desktop app]. https://www.grammarly.com/desktop/mac

Greenstone, M., & Chay, K. (2005). Does Air quality matter? Evidence from housing markets. *Journal of Political Economy*, *113*, 376–424.

Grover, S., & Singh, N. H. (2002). The quality of primary education: A case study of Madurai and Villupuram districts in Tamil Nadu, India. *Center for International Development, Harvard University*.

Grubenhoff, J. A., du Ford, K., & Roosevelt, G. E. (2007). Heat-related illness. *Clinical Pediatric Emergency Medicine*, *8*(1), 59–64.

Gupta, R., & Somanathan, E. (2023). Hot days and worker absence in the formal sector: Evidence from India. Social Science Research Network.

Hammer, J., Aiyar, Y., & Samji, S. (2007). Understanding government failure in public health services. *Economic and Political Weekly*, *42*(40), 4049-4057.

Hansen, J., Sato, M., & Ruedy, R. (2012). Perception of climate change. *Proceedings of The National Academy of Sciences*, *109*(37), E2415–E2423.

Heyes, A., & Saberian, S. (2019). Temperature and decisions: evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, *11*(2), 238–265.

Heyes, A., & Saberian, S. (2022). Hot Days, the ability to Work and climate resilience: Evidence from a representative sample of 42,152 Indian households. *Journal of Development Economics*, *155*, 102786.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of The National Academy of sciences*, *107*(35), 15367-15372.

Kenney, W. L. (1985). A review of comparative responses of men and women to heat stress. *Environmental Research*, *37*(1), 1–11.

Kjellstrom, T., Briggs, D., Freyberg, C., Lemke, B., Otto, M., & Hyatt, O. (2016). Heat, human performance, and occupational health: a key issue for the assessment of global climate change impacts. *Annual review of public health*, *37*(1), 97-112.

Kumar, P., & Khan, A. M. (2013). Human resource management in primary health care system. *Health Population Perspectives and Issues*, *36*(1-2), 66-76.

Kumar, P., Mehra, A., Inder, D., & Khan, A. M. (2014). A study of human resource policies and practices for primary health care system in Delhi. *International Journal of Medicine & Public Health*, *4*(4), 430-435.

Kremer, M., Chaudhury, N., Rogers, F. H., Muralidharan, K., & Hammer, J. (2005). Teacher absence in India: A snapshot. *Journal of The European Economic Association*, *3*(2–3), 658–667.

Lei, T., Stannard, S. R., Perry, B. G., Schlader, Z. J., Cotter, J. D., & Mündel, T. (2017). Influence of menstrual phase and arid vs. humid heat stress on autonomic and behavioural thermoregulation during exercise in trained but unacclimated women. *The Journal of Physiology*, *595*(9), 2823–2837.

Leonard, C (2018). Temperature and teacher absence: Evidence from 4,085 schools in India. University of Ottawa. Available at https://ruor.uottawa.ca/bitstreams/2609047a-78e2-40e4-88d1-229441f92068/download

Liss, A., Wu, R., Chui, K. K. H., & Naumova, E. N. (2017). Heat-related hospitalizations in older adults: An amplified effect of the first seasonal heatwave. *Scientific Reports*, *7*(1), 39581.

Makowski, K. (2009). *The daily temperature amplitude and surface solar radiation*. ETH Zurich.

O'Neill, M. S., Carter, R., Kish, J. K., Gronlund, C. J., White-Newsome, J. L., Manarolla, X., Zanobetti, A., & Schwartz, J. D. (2009). Preventing heat-related morbidity and mortality: new approaches in a changing climate. *Maturitas*, *64*(2), 98–103.

Park, J. (2017b). Will we adapt? Labor productivity and adaptation to climate change. *Discussion Paper 2017-73. Cambridge, Mass.: Harvard Environmental Economics Program*

Rao, K. D., Bhatnagar, A., & Berman, P. (2012). So many, yet few: human resources for health in India. *Human Resources for Health*, *10*, 1-9.

Ranson, M. (2014). Crime, weather, and climate change. *Journal of Environmental Economics and Management*, *67*(3), 274–302.

Rohini, P., Rajeevan, M., & Srivastava, A. K. (2016). On the variability and increasing trends of heat waves over India. *Scientific Reports*, *6*(1), 1–9.

Rudrappa, S., Agarkhed, D. V., & Vaidya, S. S. (2018). Healthcare systems: India. *Quality spine care: healthcare systems, quality reporting, and risk adjustment*, Springer, 211-224.

Saxena, N., & Rai, H. (2015). The correlation effects between recruitment, selection, training, development and employee stress, satisfaction and commitment: Findings from a survey of 30 hospitals in India. *International Journal of Healthcare Technology and Management*, *15*(2), 142-161.

Scheidt, J., Koppe, C., Rill, S., Reinel, D., Wogenstein, F., & Drescher, J. (2013). Influence of temperature changes on migraine occurrence in Germany. *International Journal of Biometeorology*, *57*, 649–654.

Schlenker, W., & Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, *5*(1), 14010.

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of The National Academy of Sciences*, *106*(37), 15594–15598.

Shayegh, S., Manoussi, V., & Dasgupta, S. (2021). Climate change and development in South Africa: the impact of rising temperatures on economic productivity and labour availability. *Climate and Development*, *13*(8), 725–735.

Singh, A., & Pallikadavath, S. (2014). Human resource shortage and inequalities at the base of India's public healthcare system. In *Anais Do Congresso Sul-Brasileiro De Medicina De Família E Comunidade* 4, 8.

Somanathan, E., Somanathan, R., Sudarshan, A., & Tewari, M. (2021). The impact of temperature on productivity and labor supply: Evidence from Indian manufacturing. *Journal of Political Economy*, *129*(6), 1797–1827.

Stern, N. (2008). The economics of climate change. *American Economic Review*, *98*(2), 1–37.

Szewczyk, W., Mongelli, I., & Ciscar, J.-C. (2021). Heat stress, labour productivity and adaptation in Europe—a regional and occupational analysis. *Environmental Research Letters*, *16*(10), 105002.

Wang, W., Xie, G., & Hao, L. (2014). Rural panel surveys in developing countries: A selective review. *Economic and Political Studies*, *2*(2), 151–177.

Yang, A. C., Fuh, J.-L., Huang, N. E., Shia, B.-C., & Wang, S.-J. (2015). Patients with migraine are right about their perception of temperature as a trigger: time series analysis of headache diary data. *The Journal of Headache and Pain*, *16*, 1–7.

Zhang, P., Deschenes, O., Meng, K., & Zhang, J. (2018). Temperature effects on productivity and factor reallocation: Evidence from a half million Chinese manufacturing plants. *Journal of Environmental Economics and Management*, *88*, 1–17.

Zhang, Y., Peng, M., Wang, L., & Yu, C. (2018). Association of diurnal temperature range with daily mortality in England and Wales: a nationwide time-series study. *Science of The Total Environment*, *619*, 291–300.

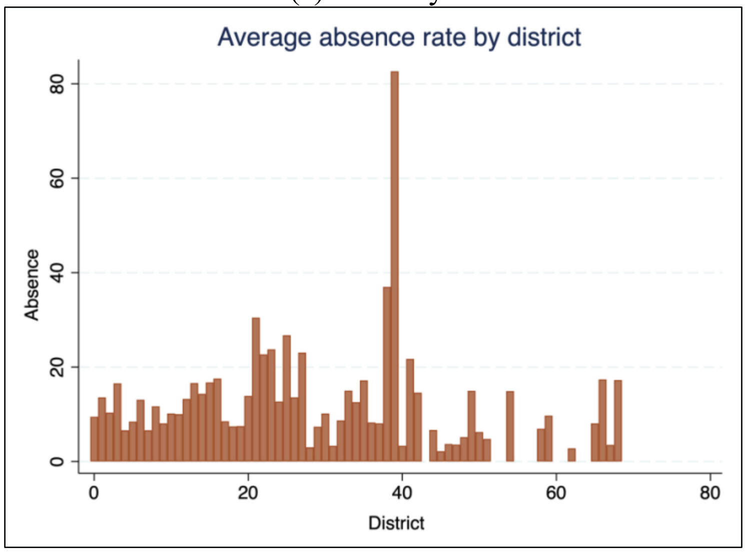
Zheng, S., Wang, M., Li, B., Wang, S., He, S., Yin, L., Shang, K., & Li, T. (2016). Gender, age and season as modifiers of the effects of diurnal temperature range on emergency room admissions for cause-specific cardiovascular disease among the elderly in Beijing. *International Journal of Environmental Research and Public Health*, *13*(5), 447.

Zivin, J. G., & Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, *102*(7), 3652–3673.

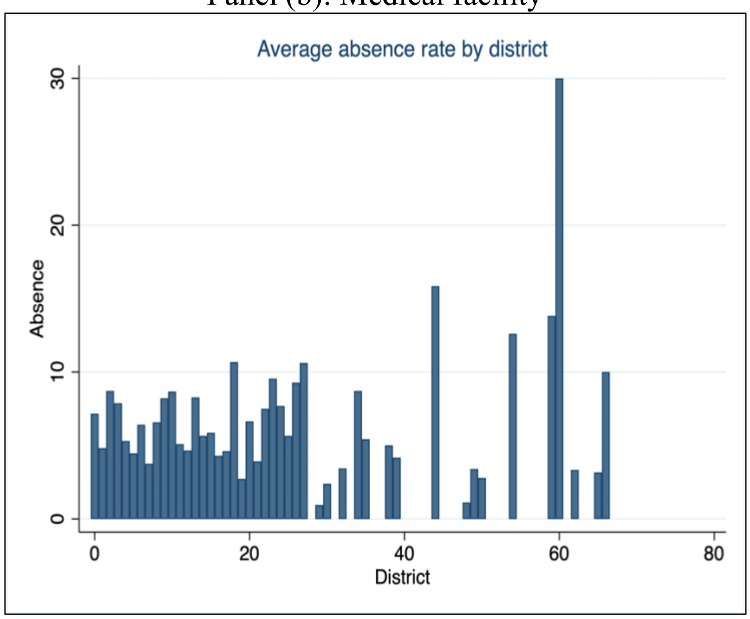
Figures and Tables

Figure 9: Distribution of absence rates

Panel (a): Primary school



Panel (b): Medical facility



Notes: Figure (9) panel (a) shows the district-wise distribution of teacher absence rate, and panel (b) shows the district-wise distribution of medical staff absence rate.

Figure 10: Temperature distribution during summer, fall, and winter

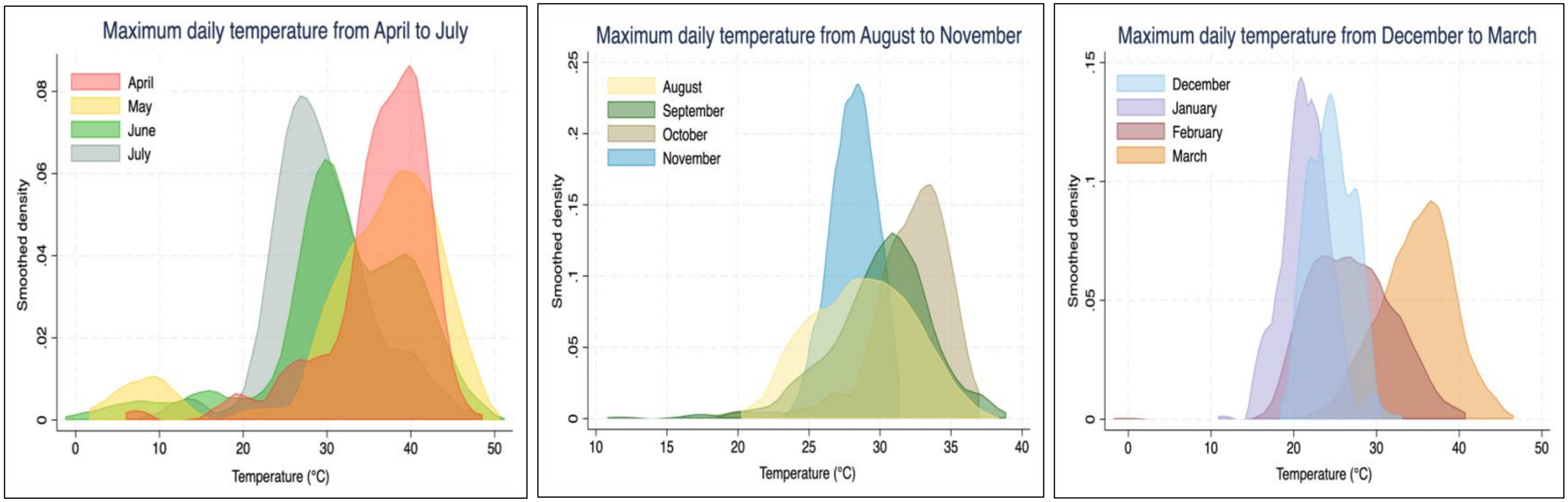
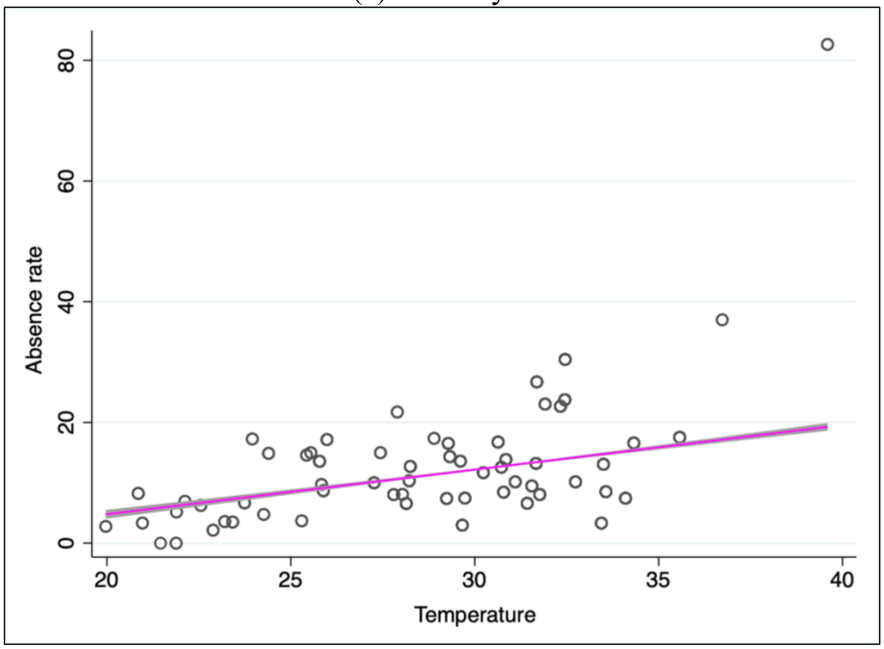


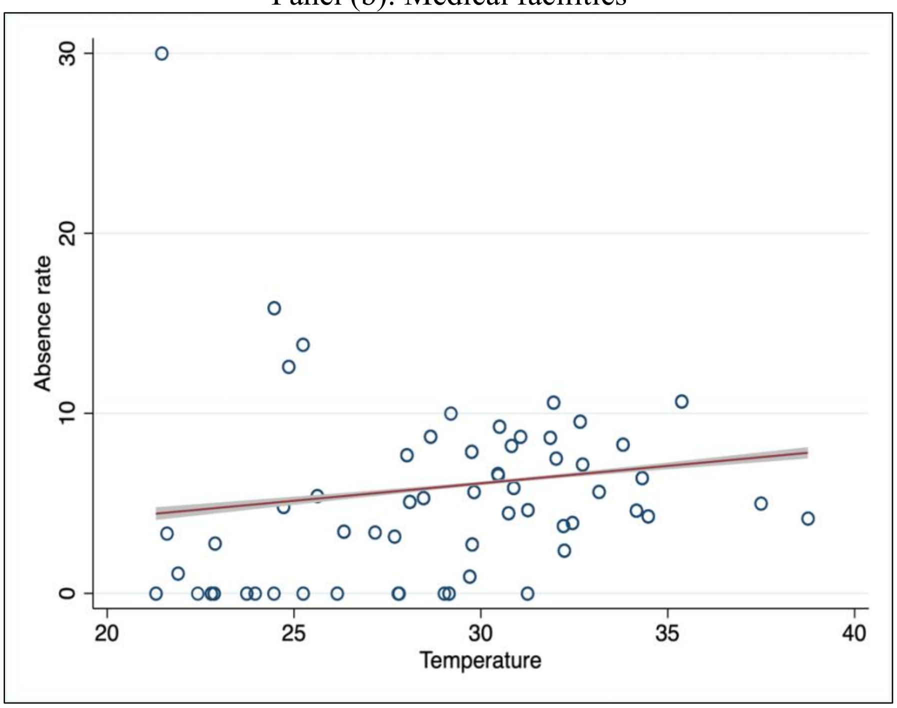
Figure (10) presents the monthly distribution of daily maximum temperature during summer, fall, and winter for the period 2011-2012.

Figure 11: Relationship between temperature and absence rates

Panel (a): Primary schools

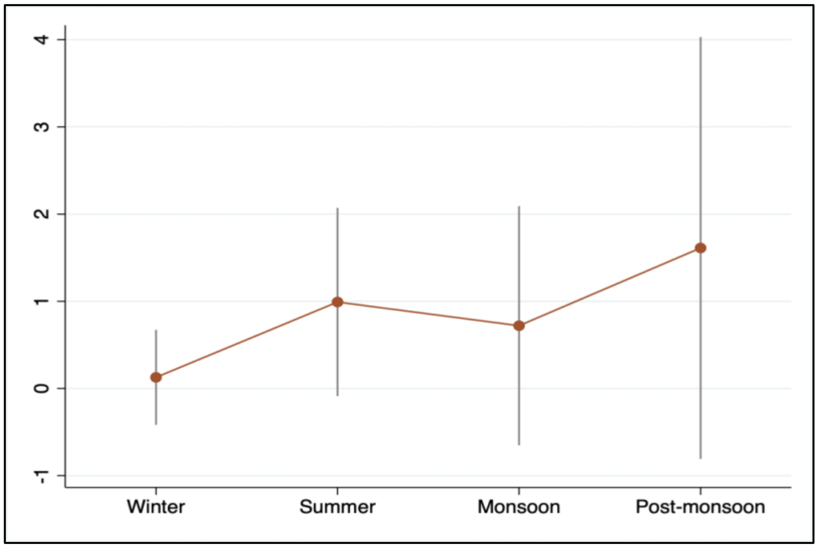


Panel (b): Medical facilities



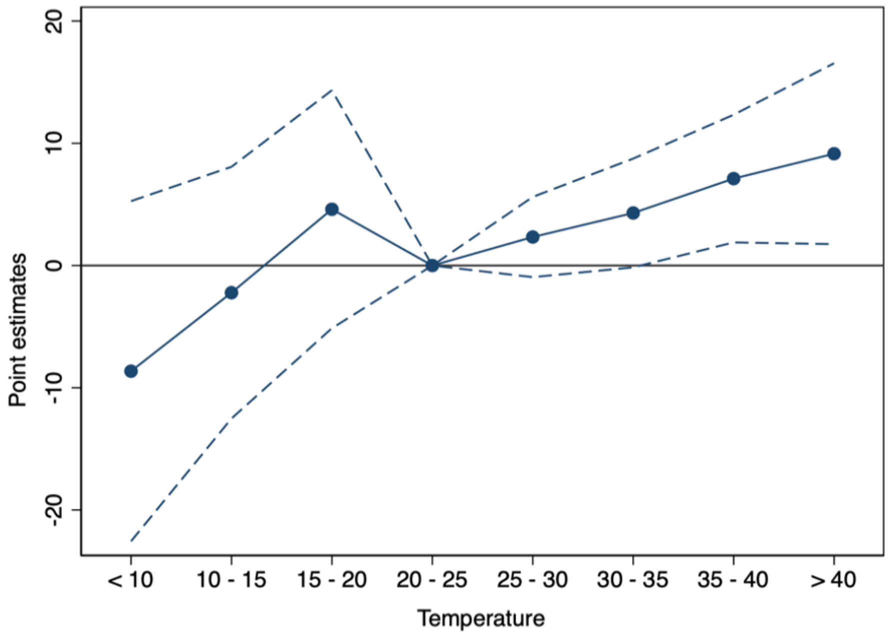
Notes: Figure (11) shows the two-way linear prediction plot of district-wise average absence rate and temperature in primary schools (panel a) and in medical facilities (panel b).

Figure 12: Estimated effect by season



Notes: Figure (12) exhibits the effect of temperature on the absence rate in four official seasons in India.

Figure 13: Non-linear estimates



Notes: Figure (13) displays a crude relationship between temperature and absence rate. Figure (5) presents the effect of temperature on absence rate considering the non-linear effect of temperature. Daily maximum temperature is divided into eight bins with a 5-degree Celsius width.

Table 19: Summary statistics

Variables	Mean	Std. Dev.
Teacher absence rate (%)	12.53	23.63
Medical staff absence rate (%)	6.23	16.54
Female teacher (%)	61.74	28.66
Male teacher (%)	38.26	28.66
Female staff (%)	48.61	30.27
Male staff (%)	51.39	30.27
Maximum daily temperature (°C)	30.47	7.40
Wind speed (km/h)	2.76	1.24
Precipitation (mm)	4.49	13.42
Solar radiation (W/m^2)	19.41	6.31
Relative humidity (%)	0.55	0.23
Electricity availability (%)	78.05	-
Water availability (%)	99.78	-

Table 20: Main results- effect of temperature on teacher absence rate

	(1) No control	(2) Weather	(3) School	(4) Day of week	(5) District	(6) Year	(7) Year-month	(8) State	(9) Preferred
Temperature	0.678^{***}	0.609^{***}	0.603^{***}	0.600^{***}	0.572^{***}	0.576^{***}	0.739^{***}	0.358^{***}	0.563^{**}
	(0.121)	(0.117)	(0.118)	(0.117)	(0.112)	(0.112)	(0.115)	(0.118)	(0.225)
Observations	4,100	4,100	4,010	4,010	4,010	4,010	4,010	4,010	4,010
Weather controls	N	Y	Y	Y	Y	Y	Y	Y	Y
School controls	N	N	Y	Y	Y	Y	Y	Y	Y
Day-of-week FE	N	N	N	Y	Y	Y	Y	Y	Y
District FE	N	N	N	N	Y	Y	Y	Y	N
State FE	N	N	N	N	N	N	N	Y	Y
Year FE	N	N	N	N	N	Y	N	N	N
Year-month FE	N	N	N	N	N	N	Y	Y	N
Year-month-district FE	N	N	N	N	N	N	N	N	Y

Notes: The dependent variable in Table 20 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Weather controls include precipitation, humidity, wind speed, and solar radiation. School controls include private-public status, school age, student enrollment, student attendance on an average day, number of classrooms, number of full-time teachers, number of part-time teachers, and number of shifts in school. Each specification adds controls indicated in the respective columns. The preferred specification includes day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Main results- effect of temperature on medical staff absence rate

	(1) No control	(2) Weather	(3) Facility	(4) Day of week	(5) District	(6) Year	(7) Year-month	(8) State	(9) Preferred
Temperature	0.054	0.071	0.155	0.148	0.108	0.101	0.065	-0.270	0.033
	(0.052)	(0.057)	(0.109)	(0.108)	(0.118)	(0.120)	(0.140)	(0.172)	(0.509)
Observations	2,747	2,747	2,747	2,747	2,747	2,747	2,747	2,747	2,747
Weather controls	N	Y	Y	Y	Y	Y	Y	Y	Y
School controls	N	N	Y	Y	Y	Y	Y	Y	Y
Day-of-week FE	N	N	N	Y	Y	Y	Y	Y	Y
District FE	N	N	N	N	Y	Y	Y	Y	N
State FE	N	N	N	N	N	N	N	Y	Y
Year FE	N	N	N	N	N	Y	N	N	N
Year-month FE	N	N	N	N	N	N	Y	Y	N
Year-month-district FE	N	N	N	N	N	N	N	N	Y

Notes: The dependent variable in Table 21 is the medical staff absence rate at each medical facility on the particular interview date. Temperature is the daily maximum temperature at the medical facility location. Weather controls include precipitation, humidity, wind speed, and solar radiation. Facility controls include private-public status, facility age, type of facility, average beds occupied at night, average outpatients each week, number of staff working at the clinic, number of sanctioned positions now vacant, and total number of operation hours per week. Each specification adds controls indicated in the respective columns. The preferred specification includes day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 22: Effect of temperature lags and leads on teacher absence rate

	(1) Preferred	(2) 1-day lag	(3) 2-day lag	(4) 1-day lead	(5) 2-day lead
Temperature	0.563**	0.578**	0.591**	0.511**	0.556**
	(0.225)	(0.234)	(0.229)	(0.239)	(0.263)
$[\text{Temperature}]_{t-1}$		-0.0693	-0.0232		
		(0.114)	(0.119)		
$[\text{Temperature}]_{t-2}$			-0.112		
			(0.157)		
$[\text{Temperature}]_{t+1}$				-0.338**	-0.326**
				(0.149)	(0.160)
$[\text{Temperature}]_{t+2}$					0.120
					(0.190)
Observations	4,010	3,665	3,324	3,666	3,325

Notes: The dependent variable in Table 22 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Column (1) represents our preferred estimation in Table 23. Columns (2) to (5) re-estimate the preferred specification with temperature lags and leads. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Effect of temperature lags and leads on medical staff absence rate

	(1) Preferred	(2) 1-day lag	(3) 2-day lag	(4) 1-day lead	(5) 2-day lead
Temperature	0.033	-0.094	0.077	-0.035	-0.071
	(0.509)	(0.544)	(0.565)	(0.518)	(0.595)
$[\text{Temperature}]_{t-1}$		0.573	0.817^{**}		
		(0.384)	(0.376)		
$[\text{Temperature}]_{t-2}$			0.205		
			(0.382)		
$[\text{Temperature}]_{t+1}$				-0.141	-0.254
				(0.450)	[0.464]
$[\text{Temperature}]_{t+2}$					-0.944^{**}
					(0.469)
Observations	2,747	991	864	987	873

Notes: The dependent variable in Table 23 is the medical staff absence rate at each medical facility on the particular interview date. Temperature is the daily maximum temperature at the medical facility location. Column (1) represents our preferred estimation in Table 23.1. Columns (2) to (5) re-estimate the preferred specification with temperature lags and leads. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 24: Gender participation rate and age at the school level

	(1) Female (>= 50)	(2) Male (>= 50)	(3) Age >=32 (25th percentile)	(4) Age >=32 & Female (>= 50)	(5) Age >=32 & male (>= 50)
Temperature	0.612^{**}	0.578	0.647^{**}	0.691^{**}	0.863
	(0.270)	(0.448)	(0.305)	(0.333)	(0.688)
Observations	2,891	1,567	2,960	2,182	1,125

Notes: The dependent variable in Table 24 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (5) re-estimate the preferred specification on male, female, and age-group subsamples. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Gender participation rate at the medical facility level

	(1) Female (>= 50)	(2) Male (>= 50)
Temperature	0.606	-0.279
	(0.809)	(0.574)
Observations	628	881

Notes: The dependent variable in Table 25 is the medical staff absence rate at each medical facility on the particular interview date. Temperature is the daily maximum temperature at the medical facility location. Columns (1) to (5) re-estimate the preferred specification on male and female subsamples. We do not have information on staff age in IHDS-II. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 26: Distance-percentage of teachers living more than a km away from school

	(1) Away (> 1 km)	(2) Away on hot days (> 29°C)	(3) Excluding top 10%	(4) Excluding top 25%	(5) Excluding top 10% on hot days (> 29°C)	(6) Excluding top 25% on hot days (> 29°C)
Temperature	0.551^{**}	0.954^{***}	0.531^{**}	0.582^{**}	0.942^{**}	1.157^{***}
	(0.224)	(0.367)	(0.230)	(0.227)	(0.368)	(0.329)
Away (percentage)	0.030^{**}	0.036^{*}	0.031^{*}	0.046^{**}	0.044^{*}	0.050
	(0.014)	(0.021)	(0.016)	(0.021)	(0.024)	(0.032)
Observations	4,010	2,322	3,524	2,988	1,999	1,669

Notes: The dependent variable in Table 26 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (6) re-estimate the preferred specification, including a measure of distance from school and excluding schools where all teachers live away. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 27: Distance-percentage of medical staff living more than a km away from the facility

	(1) Away (> 1 km)	(2) Away on hot days (> 29°C)	(3) Excluding top 10%	(4) Excluding top 25%	(5) Excluding top 10% on hot days (> 29°C)	(6) Excluding top 25% on hot days (> 29°C)
Temperature	-0.032	-0.374	0.018	-0.252	-0.529	-0.957
	(0.496)	(0.727)	(0.547)	(0.644)	(0.692)	(0.639)
Away (percentage)	0.071^{**}	0.073^{*}	0.124^{***}	0.156^{**}	0.150^{***}	0.149^{*}
	(0.029)	(0.044)	(0.037)	(0.071)	(0.047)	(0.085)
Observations	1,126	687	1,014	892	614	542

Notes: The dependent variable in Table 27 is the medical staff absence rate at each medical facility on the particular interview date. Temperature is the daily maximum temperature at the medical facility location. Columns (1) to (6) re-estimate the preferred specification, including a measure of distance from the facility and excluding facilities where all staff live away. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 28: Access to electricity

	(1) Interaction	(2) Electricity control	(3) 1-4 hours	(4) 5-8 hours
Temperature	0.600^{***}	0.555^{**}	0.799	0.503^{*}
	(0.231)	(0.226)	(0.542)	(0.299)
Electricity*Temperature	-0.055			
	(0.037)			
Electricity		-1.613		
		(1.100)		
Observations	4,002	4,002	1,097	2,042

Notes: The dependent variable in Table 28 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (4) re-estimate the preferred specification for schools with electricity controls, length of electricity supply, and an interaction of temperature and electricity access. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 29: Quality of power supply

	(1) Interaction	(2) Outage control	(3) Never or less frequent	(4) Everyday
Temperature	0.502*	0.515*	0.382	0.653*
	(0.277)	(0.280)	(0.585)	(0.365)
Outage* Temperature	0.013			
	(0.048)			
Outage		-0.419		
		(1.375)		
Observations	3,059	3,059	1,092	1,967

Notes: The dependent variable in Table 29 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (4) re-estimate the preferred specification, including a control for power outages and the frequency of power outages. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 30: Cooling accommodation

	(1) Interaction	(2) Fan control	(4) Fan	(5) No Fan
Temperature	0.568**	0.565**	0.391	0.778
	(0.226)	(0.225)	(0.309)	(0.479)
Fan*Temperature	-0.004			
	(0.034)			
Fan		-0.201		
		(0.982)		
Observations	4,007	4,007	2,760	1,247

Notes: The dependent variable in Table 30 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (5) re-estimate the preferred specification for schools with and without fans, including fan control and the interaction of temperature and fan ownership. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 31: Alternative measures of temperature

	(1) Mean > 25th percentile	(2) Minimum < 25th percentile	(3) DTR	(4) DTR < 75th percentile	(5) Interaction	(6) Solar radiation < 75th percentile	(7) Heat index
Average temperature	0.874**						
	(0.387)						
Minimum temperature		-0.888**					
		(0.420)					
DTR			0.858***	1.030***		1.115***	
			(0.329)	(0.381)		(0.374)	
DTR*humidity					0.699*		
					(0.407)		
Heat index							0.126*
							(0.069)
Observations	2,823	1,009	4,010	2,927	4,010	3,015	4,010

Notes: The dependent variable in Table 31 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (6) re-estimate the preferred specification, replacing temperature with daily average and minimum temperature, and different specifications for the diurnal temperature range. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 32: Indoor vs outdoor schools

	(1) Interaction	(2) Indoor control	(3) Indoor schools
Temperature	0.545**	0.545**	0.646***
	(0.235)	(0.224)	(0.230)
Indoor*Temperature	0.000		
	(0.057)		
Indoor		-0.0304	
		(1.545)	
Observations	3,981	3,981	3,628

Notes: The dependent variable in Table 32 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (5) re-estimate the preferred specification for schools with indoor classrooms, including indoor control and interaction of temperature and indoor classroom status. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 33: Grant status

	(1) With grant	(2) No grant
Temperature	0.500	1.012**
	(0.385)	(0.446)
Observations	2,102	990

Notes: The dependent variable in Table 33 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (5) re-estimate the preferred specification for schools with and without grants. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 34: Evaluation criteria

	(1) No	(2) Yes	(3) Non-attendance criteria
Temperature	0.455	0.673**	0.774**
	(0.341)	(0.326)	(0.327)
Observations	1,529	2,404	2,602

Notes: The dependent variable in Table 34 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (5) re-estimate the preferred specification for schools with and without performance evaluation criteria for teachers. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 35: Alternative standard errors

	(1) District preferred	(2) School	(3) School-month	(4) District-year	(5) State-year	(6) Year-month	(7) Climate zone-month	(8) Eicker-White
Temperature	0.563**	0.563***	0.563***	0.563**	0.563**	0.563***	0.563***	0.563***
	(0.225)	(0.211)	(0.211)	(0.224)	(0.254)	(0.198)	(0.202)	(0.211)
Observations	4,010	4,010	4,010	4,010	4,010	4,010	4,010	4,010

Notes: The dependent variable in Table 35 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Column (1) represents our preferred specification in column (8) of Table 2. Columns (2) to (5) re-estimate the preferred specification by clustering the standard errors at district-year, state-year, year-month, and climate zone-month levels. Column (6) considers Eiker-White standard errors. Standard errors are in brackets. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 36: Effect by climate zones

	(1) Peninsular	(2) West central	(3) North west	(4) North central	(5) North east	(6) Northern hilly
Temperature	0.145	0.977	1.185***	0.172	0.305	0.697
	(0.698)	(1.039)	(0.445)	(0.491)	(0.394)	(0.573)
Observations	910	790	840	848	350	258

Notes: The dependent variable in Table 36 is the teacher absence rate at each school on the particular interview date. Temperature is the daily maximum temperature at the school location. Columns (1) to (6) re-estimate the preferred specification at six different climate zones. All regressions include day-of-week, year, year-month, year-month-district, district, and state dummies. Standard errors in brackets are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Conclusion

Environmental degradation caused by climate change poses significant challenges to human capital development and economic growth. This dissertation investigates how extreme environmental shocks impact human capital stock at different stages of life. Using nationally representative datasets and econometric methods, this research provides causal evidence regarding the effects of air pollution, flooding, and rising temperatures on child mortality, children's cognitive abilities, and employee absence at work. The findings underscore the importance of preventive and mitigative policies in alleviating the impacts of environmental shocks on human health and well-being. Furthermore, it unfolds several new research topics for further exploration, including comparisons of existing and new policies to determine more effective ways to mitigate environmental pollution, assessing the long-term effects of environmental shocks across generations, and examining geographic variations in these effects.

Appendix

Table A1: OLS effects of $PM_{2.5}$ on child mortality

	(1) Early neonatal	(2) Neonatal	(3) Post- neonatal	(4) Infant	(5) One-to-five	(6) Under five
Mean daily PM2.5	-0.026	-0.029	-0.022	-0.052*	-0.000	-0.010*
	(0.019)	(0.022)	(0.014)	(0.029)	(0.001)	(0.005)
Temperature						
Bin 2	-4.817	-6.081	-1.310	-7.286	20.418**	14.504*
	(7.044)	(7.236)	(8.031)	(12.354)	(9.037)	(8.433)
Bin 3	8.767	7.369	3.425	10.542	12.069**	11.774*
	(10.508)	(10.661)	(8.724)	(16.175)	(4.659)	(6.086)
Bin 4	21.757	25.509*	31.985	57.298*	9.565**	19.919**
	(12.373)	(13.727)	(23.012)	(30.004)	(4.122)	(9.483)
Bin 5	35.036**	37.662***	27.019	64.534**	11.862*	23.443**
	(10.713)	(10.287)	(20.326)	(24.411)	(6.074)	(9.871)
Bin 6	31.586**	32.267**	14.721*	46.882**	11.180**	18.724**
	(10.115)	(10.104)	(8.160)	(16.290)	(5.027)	(7.018)
Bin 7	28.889***	33.302***	10.167	43.488***	10.474**	17.436**
	(6.805)	(6.894)	(7.684)	(12.149)	(4.197)	(5.514)
Bin 8	29.071***	31.864***	12.332	44.235***	11.023**	18.088**
	(7.602)	(7.529)	(7.593)	(12.888)	(4.419)	(5.816)
Bin 9	24.808***	28.195***	12.691*	40.962**	10.641**	17.109**
	(6.861)	(6.972)	(7.467)	(12.611)	(4.225)	(5.563)
Bin 10	27.643***	30.689***	12.871*	43.663***	10.558**	17.542**
	(6.749)	(6.822)	(7.667)	(12.438)	(4.210)	(5.503)
Bin 11	26.221***	31.271***	14.112*	45.545***	10.942**	18.238**
	(6.771)	(6.851)	(7.640)	(12.159)	(4.199)	(5.433)
Bin 12	26.218***	30.094***	13.244*	43.433***	11.163**	18.025**
	(6.804)	(6.907)	(7.581)	(12.419)	(4.201)	(5.460)
Bin 13	27.219***	31.461***	12.152	43.631***	10.776**	17.742**
	(6.867)	(7.068)	(7.561)	(12.448)	(4.217)	(5.441)

Bin 14	27.452***	32.286***	13.034*	45.357***	10.864**	18.196**
	(6.713)	(7.012)	(7.519)	(12.350)	(4.244)	(5.460)
Bin 15	27.660***	32.285***	14.270*	46.540***	10.326**	17.972**
	(6.915)	(7.157)	(7.617)	(12.467)	(4.216)	(5.459)
Bin 16	27.662***	31.924***	14.011*	45.875***	10.706**	18.153**
	(6.869)	(7.046)	(7.503)	(12.394)	(4.208)	(5.445)
Bin 17	29.488***	34.247***	13.254*	47.511***	10.871**	18.599***
	(7.122)	(7.466)	(7.727)	(12.613)	(4.196)	(5.450)
Bin 18	31.016***	36.043***	14.158*	50.325***	10.871**	19.139***
	(7.094)	(7.409)	(7.610)	(12.579)	(4.336)	(5.543)
Bin 19	29.955***	35.629***	14.343*	50.054***	10.696**	18.948***
	(7.014)	(7.303)	(7.528)	(12.359)	(4.289)	(5.475)
Bin 20	30.604***	35.626***	15.475**	51.199***	10.769**	19.255***
	(7.324)	(7.566)	(7.402)	(12.501)	(4.354)	(5.521)
Bin 21	31.842***	38.220***	15.146*	53.480***	10.744**	19.662***
	(7.438)	(7.632)	(7.761)	(12.836)	(4.317)	(5.538)
Bin 22	27.207***	33.450***	12.507	45.951***	10.977**	18.415**
	(7.503)	(7.831)	(7.726)	(12.739)	(4.345)	(5.554)
Bin 23	31.288***	36.638***	14.916*	51.520***	11.229**	19.666***
	(7.804)	(8.124)	(8.027)	(13.019)	(4.390)	(5.589)
Bin 24	58.821**	60.897**	13.620	74.030**	11.962**	25.056***
	(24.864)	(24.585)	(9.272)	(26.886)	(4.537)	(7.192)
Bin 25	17.837*	30.164**	6.118	35.576**	9.480**	15.328**
	(8.924)	(12.627)	(8.326)	(15.999)	(4.372)	(5.821)
Wind speed	-0.077	-0.109	-0.035	-0.147*	0.007	-0.023
	(0.063)	(0.068)	(0.036)	(0.080)	(0.008)	(0.017)
Precipitation	-0.528	0.157	0.727	0.872	-0.049	0.151
	(0.867)	(0.936)	(0.599)	(1.149)	(0.090)	(0.245)
Humidity	0.050	0.052	-0.014	0.036	-0.004	0.002
	(0.032)	(0.037)	(0.027)	(0.049)	(0.005)	(0.011)
Cloud cover	-0.014	-0.023	0.007	-0.014	0.001	-0.001
	(0.021)	(0.024)	(0.015)	(0.030)	(0.002)	(0.007)
Observations	500,310	500,310	500,310	500,310	500,310	500,310

Table A2: IV effects of $PM_{2.5}$ on child mortality

	(1) Early neonatal	(2) Neonatal	(3) Post- neonatal	(4) Infant	(5) One-to-five	(6) Under five
Mean daily PM2.5	0.951**	1.029**	0.019	1.025**	-0.033	0.174**
	(0.343)	(0.355)	(0.209)	(0.410)	(0.035)	(0.083)
Temperature						
Bin 2	-3.535	-5.904	1.02	-4.436	19.499**	14.300*
	(9.540)	(9.264)	(7.221)	(14.688)	(8.856)	(8.090)
Bin 3	10.402	8.206	5.085	13.459	11.262**	11.655*
	(12.603)	(12.262)	(8.143)	(18.754)	(3.807)	(6.135)
Bin 4	22.731	25.506*	33.813	59.540*	8.817**	19.755**
	(13.156)	(14.303)	(23.737)	(33.399)	(3.175)	(9.578)
Bin 5	36.233**	38.134***	28.946	67.340**	11.146**	23.412**
	(11.764)	(10.935)	(20.887)	(28.115)	(5.137)	(10.072)
Bin 6	32.681**	32.571**	16.451**	49.326**	10.439**	18.605**
	(13.097)	(12.630)	(7.693)	(19.300)	(4.097)	(7.155)
Bin 7	30.254***	33.847***	12.060*	46.307**	9.736**	17.392**
	(8.939)	(8.611)	(6.813)	(14.417)	(3.269)	(5.571)
Bin 8	30.179**	32.205***	14.223**	46.851**	10.315**	18.022**
	(10.059)	(9.619)	(6.690)	(15.393)	(3.488)	(5.915)
Bin 9	26.183**	28.815***	14.562**	43.830**	9.934**	17.094**
	(8.971)	(8.633)	(6.627)	(14.657)	(3.290)	(5.605)
Bin 10	29.114**	31.436***	14.829**	46.740**	9.827**	17.547**
	(8.891)	(8.424)	(6.760)	(14.394)	(3.257)	(5.524)
Bin 11	27.549**	31.808***	16.103**	48.444***	10.250**	18.237***
	(8.803)	(8.365)	(6.695)	(14.077)	(3.244)	(5.452)
Bin 12	27.340**	30.376***	15.201**	46.043**	10.471**	17.968**
	(8.961)	(8.529)	(6.535)	(14.227)	(3.243)	(5.460)
Bin 13	28.283**	31.622***	14.149 **	46.161**	10.079**	17.664**
	(8.736)	(8.420)	(6.551)	(14.021)	(3.260)	(5.417)
Bin 14	28.325**	32.214***	15.056**	47.684***	10.171**	18.077***
	(8.587)	(8.349)	(6.500)	(13.853)	(3.282)	(5.413)

Bin 15	28.249**	31.809***	16.352**	48.525***	9.661**	17.810**
	(8.776)	(8.499)	(6.533)	(13.967)	(3.261)	(5.416)
Bin 16	27.812**	30.997***	16.106**	47.435***	10.058**	17.920***
	(8.681)	(8.349)	(6.495)	(13.852)	(3.255)	(5.383)
Bin 17	29.004**	32.593***	15.376**	48.385***	10.243**	18.247***
	(8.899)	(8.703)	(6.576)	(13.983)	(3.242)	(5.374)
Bin 18	29.636**	33.352***	16.305**	50.201***	10.288**	18.626***
	(8.838)	(8.686)	(6.547)	(13.937)	(3.397)	(5.468)
Bin 19	27.289**	31.526***	16.412**	48.470***	10.157**	18.183***
	(8.712)	(8.552)	(6.606)	(13.635)	(3.363)	(5.372)
Bin 20	26.604**	30.070***	17.611**	48.257***	10.281**	18.264***
	(8.880)	(8.766)	(6.494)	(13.685)	(3.446)	(5.409)
Bin 21	25.634**	30.255***	17.316**	48.213***	10.338**	18.276***
	(9.022)	(8.938)	(6.884)	(13.948)	(3.441)	(5.405)
Bin 22	18.239*	22.498**	14.734**	37.813**	10.677**	16.551**
	(9.304)	(9.347)	(7.001)	(13.901)	(3.504)	(5.408)
Bin 23	19.908*	23.036**	17.099**	40.737**	10.950**	17.299**
	(9.893)	(9.978)	(7.370)	(14.274)	(3.596)	(5.440)
Bin 24	45.347*	45.004*	15.944*	61.176**	11.742**	22.322**
	(25.834)	(25.391)	(8.715)	(27.785)	(3.796)	(7.129)
Bin 25	3.567	12.263	8.633	20.776	9.240**	12.223**
	(10.808)	(13.933)	(7.518)	(17.019)	(3.568)	(5.656)
Wind speed	0.069	0.050	-0.026	0.017	-0.001	0.001
	(0.078)	(0.080)	(0.048)	(0.091)	(0.010)	(0.020)
Precipitation	-0.036	0.675	0.715	1.365	-0.058	0.241
	(0.930)	(1.003)	(0.614)	(1.237)	(0.095)	(0.261)
Humidity	0.015	0.012	-0.018	-0.006	-0.003	-0.005
	(0.032)	(0.038)	(0.029)	(0.053)	(0.005)	(0.011)
Cloud cover	-0.023	-0.031	0.006	-0.023	0.001	-0.003
	(0.021)	(0.025)	(0.014)	(0.029)	(0.002)	(0.006)
Observations	500,310	500,310	500,310	500,310	500,310	500,310

Figure A1: Monthly $PM_{2.5}$ concentration in Essex, Ontario

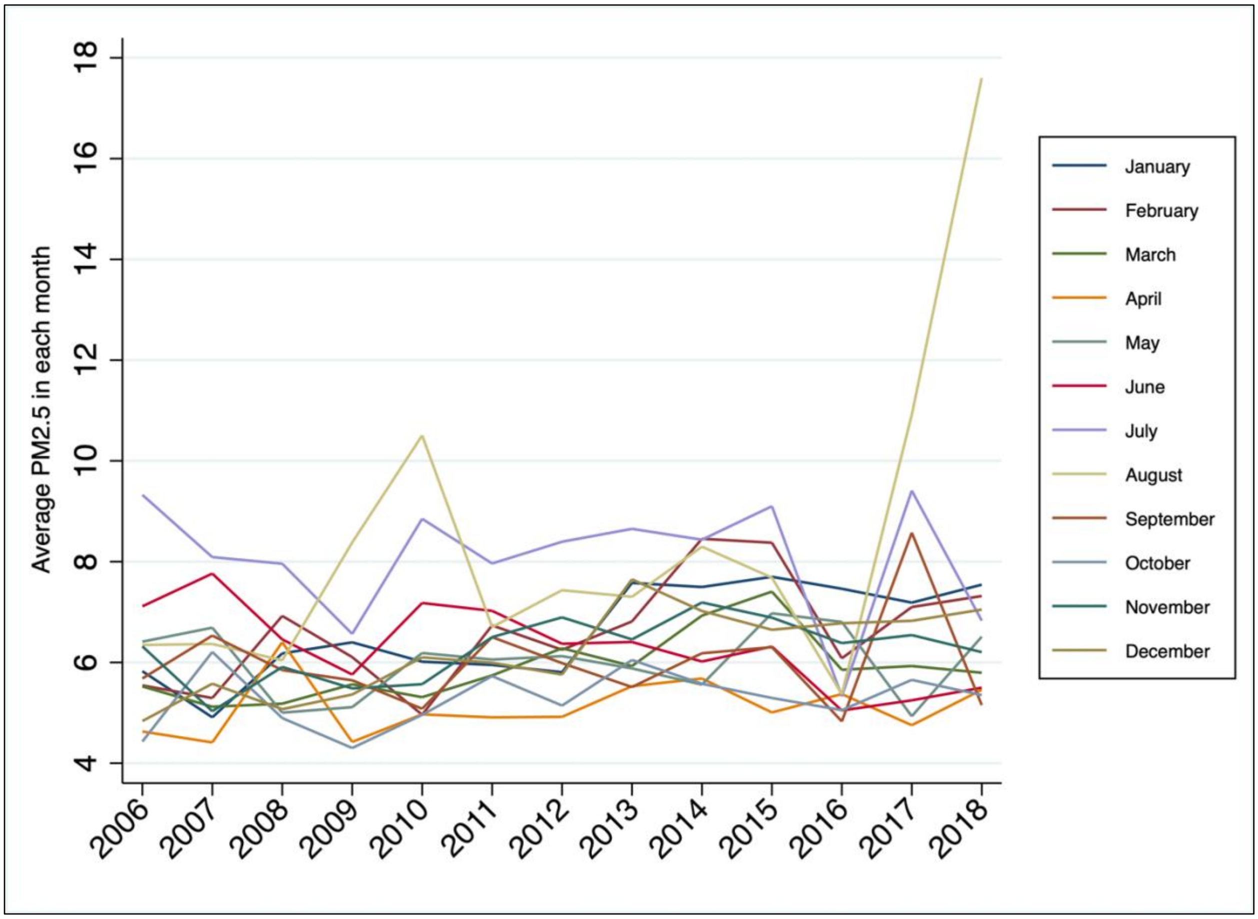


Table A3: Summary of test scores by cycle

Cycle	Test	Mean
1	Math	405.265
	PPVT	121.005
	Motor and social development	99.855
2	Math	441.225
	PPVT	97.621
	Motor and social development	99.562
	Math	846.49
Before flood	PPVT	218.626
	Motor and social development	199.417
3	Math	430.608
	PPVT	99.378
	Motor and social development	99.639
4	Math	425.308
	PPVT	100.078
	Motor and social development	98.807
5	Math	428.394
	PPVT	99.563
	Motor and social development	100.044
6	Math	470.257
	PPVT	100.959
	Motor and social development	101.051
7	Math	409.906
	PPVT	98.768
	Motor and social development	100.547
8	Math	455.337
	PPVT	97.662
	Motor and social development	99.699
	Math	2,619.81
After flood	PPVT	596.408
	Motor and social development	599.787
Total	Math	431.039
	PPVT	100.914
	Motor and social development	99.857

Table A4: Effect of flood on test scores

	Mean score
Flood × Distance	0.385***
	(0.125)
Age	32.62***
	(0.310)
Male	-52.05
	(46.48)
Parent 1 education- Secondary school graduation	-19.87***
	(4.034)
Parent 1 education- Beyond high school	-25.03***
	(4.095)
Parent 1 education- College or university degree	-18.56***
	(4.393)
Parent 1 education- Other	-29.97*
	(16.72)
Parent 1 education- Not stated	-1.501
	(8.330)
Parent 2 education- Secondary school graduation	-10.79*
	(5.772)
Parent 2 education- Beyond high school	-17.73***
	(5.902)
Parent 2 education- College or university degree	-9.365*
	(5.380)
Parent 2 education- Other	13.88
	(15.69)
Parent 2 education- Not stated	-10.94**
	(4.790)
Recoded household income- 10,000 to 14,999	9.987
	(7.363)
Recoded household income- 15,000 to 19,999	10.04
	(6.296)
Recoded household income- 20,000 to 29,999	3.867
	(7.434)
Recoded household income- 30,000 to 39,999	6.155
	(7.539)
Recoded household income- 40,000 or more	4.008
	(7.797)
Pre-maturity- Child born in normal range - gestational age 259-293 days	6.208
	(4.948)
Pre-maturity- Child born late - gestational age 294 days or late	-16.22
	(12.02)
Pre-maturity-Not stated	-62.51***
	(4.733)
Observations	11,310

Table A5: Effect of temperature on teacher absence rate

	(1) No control	(2) Weather	(3) School	(4) Day of week	(5) District	(6) Year	(7) Year-month	(8) State	(9) Preferred
Temperature	0.678***	0.609***	0.603***	0.600***	0.572***	0.576***	0.739***	0.358***	0.563**
	(0.121)	(0.117)	(0.118)	(0.117)	(0.112)	(0.112)	(0.115)	(0.118)	(0.225)
Precipitation		0.032	0.043	0.043	0.026	0.027	0.008	0.019	0.045
		(0.040)	(0.040)	(0.039)	(0.043)	(0.043)	(0.041)	(0.046)	(0.060)
Wind speed		1.501**	1.659**	1.672**	1.264**	1.271**	-0.201	0.225	-0.191
		(0.699)	(0.673)	(0.672)	(0.569)	(0.584)	(0.559)	(0.603)	(0.777)
Relative humidity		-5.803	-5.126	1.672**	-7.346**	-7.362**	-3.815	-9.223*	-2.539
		(4.079)	(3.887)	(0.672)	(3.552)	(3.644)	(5.361)	(5.143)	(5.801)
Solar radiation		-0.016	-0.014	-0.016	-0.059	-0.056	-0.237*	-0.010	-0.009
		(0.150)	(0.140)	(0.139)	(0.120)	(0.123)	(0.122)	(0.130)	(0.153)
School status (private)			-2.989***	-2.999***	-3.038***	-3.039***	-3.142***	-3.593***	-2.566***
			(1.124)	(1.124)	(1.042)	(1.043)	(1.013)	(1.019)	(0.987)
Age of school			-0.059***	-0.058***	-0.053***	-0.053***	-0.052***	-0.051***	-0.039**
			(0.019)	(0.019)	(0.018)	(0.018)	(0.017)	(0.017)	(0.016)
Student enrollment			-0.002	-0.002	-0.004	-0.004	-0.004	-0.004	-0.005
			(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.009)
Student attendance			0.004	0.004	0.004	0.004	0.003	0.003	0.006
			(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
Number of classrooms			-0.048	-0.051	0.030	0.029	-0.028	-0.076	-0.229
			(0.174)	(0.175)	(0.178)	(0.177)	(0.171)	(0.181)	(0.245)
Number of full-time teachers			-0.209	-0.203	-0.171	-0.169	-0.115	-0.009	-0.016
			(0.156)	(0.154)	(0.141)	(0.141)	(0.134)	(0.142)	(0.148)
Number of part-time teachers			0.624*	0.619*	0.809**	0.805**	0.916**	0.958**	0.911**
			(0.372)	(0.373)	(0.376)	(0.376)	(0.370)	(0.371)	(0.420)
Number of shifts			5.332*	5.246*	3.829	3.831	3.776*	2.902	2.203
			(3.017)	(2.986)	(2.335)	(2.337)	(2.085)	(2.134)	(2.600)
Observations	4,100	4,100	4,010	4,010	4,010	4,010	4,010	4,010	4,010
Day-of-week FE	N	N	Y	N	Y	Y	Y	Y	Y
District FE	N	N	Y	N	N	Y	Y	Y	N
State FE	N	N	N	N	N	N	N	Y	Y
Year FE	N	N	N	N	N	Y	N	N	N
Year-month FE	N	N	N	N	N	N	Y	Y	N
Year-month-district FE	N	N	N	N	N	N	N	N	Y

Table A6: Effect of temperature on medical staff absence rate

	(1) No control	(2) Weather	(3) Facility	(4) Day of week	(5) District	(6) Year	(7) Year-month	(8) State	(9) Preferred
Temperature	0.054	0.071	0.155	0.148	0.108	0.101	0.065	-0.270	0.033
	(0.052)	(0.057)	(0.109)	(0.108)	(0.118)	(0.120)	(0.140)	(0.172)	(0.509)
Precipitation		0.050	0.095	0.093	0.0766	0.074	0.061	0.084	0.033
		(0.042)	(0.088)	(0.088)	(0.089)	(0.090)	(0.092)	(0.091)	(0.131)
Wind speed		0.236	0.713	0.697	0.592	0.555	0.088	0.159	1.308
		(0.330)	(0.532)	(0.539)	(0.559)	(0.563)	(0.649)	(0.673)	(1.126)
Relative humidity		-1.930	-0.879	-0.477	-0.441	-0.714	1.457	-8.423	11.54
		(2.194)	(4.039)	(4.076)	(4.696)	(4.797)	(6.243)	(6.379)	(18.40)
Solar radiation		-0.083	-0.148	-0.130	-0.146	-0.156	-0.215	0.001	-0.032
		(0.071)	(0.142)	(0.143)	(0.122)	(0.127)	(0.156)	(0.157)	(0.230)
Facility status (private)			-13.95***	-13.98***	-13.64***	-13.60***	-13.62***	-13.19***	-15.66***
			(1.673)	(1.692)	(1.904)	(1.908)	(1.949)	(2.075)	(3.172)
Age of facility			-0.040*	-0.040*	-0.036	-0.036	-0.036	-0.027	-0.058
			(0.022)	(0.023)	(0.026)	(0.025)	(0.027)	(0.025)	(0.039)
Facility type			0.173	0.183	0.143	0.139	0.166	0.085	0.170
			(0.323)	(0.325)	(0.354)	(0.355)	(0.365)	(0.376)	(0.576)
Average beds occupied at night			-0.007	-0.007	-0.007	-0.007	-0.005	-0.004	0.005
			(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)
Average out-patients each week			0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Number of staff working at clinic			-0.024**	-0.023*	-0.021*	-0.021*	-0.021*	-0.028***	-0.020*
			(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.010)	(0.011)
Number of sanctioned positions now vacant			-0.012	-0.012	0.005	0.005	0.005	0.002	-0.020
			(0.033)	(0.035)	(0.035)	(0.035)	(0.033)	(0.032)	(0.048)
Total hours of operation per week			-0.032*	-0.031*	-0.041*	-0.041*	-0.044*	-0.067***	-0.074**
			(0.018)	(0.018)	(0.024)	(0.024)	(0.025)	(0.023)	(0.032)
Observations	2,747	2,747	1,126	1,126	1,126	1,126	1,126	1,126	1,126
Day-of-week FE	N	N	Y	N	Y	Y	Y	Y	Y
District FE	N	N	Y	N	N	Y	Y	Y	N
State FE	N	N	N	N	N	N	N	Y	Y
Year FE	N	N	N	N	N	Y	N	N	N
Year-month FE	N	N	N	N	N	N	Y	Y	N
Year-month-district FE	N	N	N	N	N	N	N	N	Y