UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Essays on Health and the Environment in Brazil

Ensaios em Saúde e Meio Ambiente no Brasil

Tales Rozenfeld

Orientador: Prof. Dr. Ariaster Baumgratz Chimeli

São Paulo
2021

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. Fábio Frezatti
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. José Carlos de Souza Santos
Chefe do Departamento de Economia

Prof. Dr. Ariaster Baumgratz Chimeli
Coordenador do Programa de Pós-Graduação em Economia

TALES ROZENFELD

Tese apresentada ao Departamento de
Economia da Faculdade de Economia,
Administração e Contabilidade da Uni-
versidade de São Paulo como requisito
parcial para a obtenção do título de
Doutor em Ciências.

Área de concentração: Teoria Econômica

Orientador: Prof. Dr. Ariaster Baumgratz Chimeli

Versão Corrigida

São Paulo
2021

AGRADECIMENTOS

Começo esta seção agradecendo à Bruna Borges, minha colega de doutorado e companheira de vida. Agradeço todo o suporte emocional, técnico, e, por que não, burocrático. Sem ela estas linhas ficariam incompletas. Eu te amo e muito obrigado por tudo.

Gostaria também de deixar meu agradecimento à minha família por todo suporte ao longo desse caminho. Sempre recebi todo incentivo e confiança que necessitei e sou muito grato por isso. Obrigado Cibele, Henrique, Tatiana e Therezinha pelo apoio incondicional. Obrigado Marcia e João Fernando por abrirem as portas do apartamento em São Paulo nos momentos de maior sufoco dessa caminhada e pelas horas de descontração durante o fatídico isolamento social. Obrigado Sandra, Waldemar e Carolina pelo suporte em todas as nossas decisões. Diversos amigos e colegas também foram fundamentais nesta jornada. Deixo um agradecimento a todos que tornaram esses anos mais leves e divertidos.

Gostaria de agradecer muito ao professor Ariaster Chimeli pela paciência com a qual me orientou e pela liberdade que me deu no desenvolvimento desta pesquisa. Aprendi muito nestes anos de convivência, e, dentre os principais aprendizados, destaco o estímulo à interação e cooperação entre pesquisadores. Neste sentido, agradeço também a Keyi, Luiza, Liz, Priscila, Victor e Rafael, colegas de grupo de estudos que colaboraram com o desenvolvimento deste trabalho.

Sou muito grato aos professores Rodrigo Soares e Wolfram Schlenker. Pela boa vontade e paciência deles, tive uma das vivências acadêmicas mais interessantes da minha vida na Universidade de Columbia. Eles me receberam de braços abertos, fazendo excelentes contribuições para esta pesquisa, e facilitando minha inserção na vida universitária de Columbia. Agradeço, também, aos professores André Chagas, Danilo Igliori, Paula Pereda, Marcos Nakaguma, Eduardo Haddad, e Carlos Roberto Azzoni por suas sugestões e contribuições para este trabalho.

também aos diversos funcionários da FEA por todo suporte ao longo desses anos e
à Fundação Instituto de Pesquisas Econômicas.

RESUMO

Cientistas acreditam que, com as mudanças climáticas, algumas regiões da terra irão observar um aumento na frequência e na intensidade de secas. Esta tese de doutorado consiste em dois estudos que exploram os possíveis problemas de saúde que podem surgir com um aumento da escassez hídrica. No primeiro estudo, investigamos os efeitos de medidas implementadas repentinamente pela companhia de água da Região Metropolitana de São Paulo para lidar com uma seca inesperada que atingiu a região em 2014 e 2015. Apesar de a infraestrutura de água e saneamento da região ser robusta, a seca foi tão intensa que levou o sistema para a beira do colapso, e a empresa de água precisou implementar uma série de medidas restritivas para garantir a continuidade do abastecimento. Uma importante medida desse período foi a redução da pressão da água na rede de distribuição para evitar vazamentos por pontos de fratura da tubulação. Nossa análise sugere que, apesar de importante, essa política teve o resultado inesperado de aumentar o número de hospitalizações por doenças de transmissão feco-oral. Alinhado às conclusões da literatura, nossos resultados sugerem que os grupos mais afetados são crianças e idosos de baixa renda. Analisando heterogeneidades espaciais na variação do preço de garrafas de água mineral, encontramos evidências de que a população adotou um comportamento defensivo para se proteger dos problemas relacionados ao abastecimento de água. Essas ações defensivas parecem ter afetado os indicadores de saúde observados na região. No segundo estudo desta tese, discutimos uma outra consequência da escassez de água que surge em regiões que, assim como o Brasil, tem grande dependência de usinas hidrelétricas. Nesses lugares, uma seca pode impactar a matriz energética levando a uma maior dependência de usinas termelétricas. Evidências sugerem que esse tipo de energia pode afetar a saúde da população, e termelétricas movidas a carvão costumam ocupar o centro das discussões pela sua elevada taxa de emissão. Apesar disso, outros combustíveis de termelétricas também podem influenciar indicadores de saúde de comunidades vizinhas. Neste trabalho, avaliamos o impacto de diferentes tipos de termelétricas na saúde da população. Nossos resultados sugerem que as termelétricas a carvão são as mais prejudiciais à saúde, levando a um aumento da mortalidade infantil em regiões vizinhas. Os resultados também são robustos em sugerir que meninas são

mais afetadas que meninos. Nesta análise, encontramos ainda evidência de seleção fetal, o que sugere que os resultados observados estão subestimados.

Palavras-chaves: Meio Ambiente e Desenvolvimento, Saúde, Políticas de Oferta de Água, Geração Termelétrica

ABSTRACT

As climate changes, scientists expect that some parts of the earth will experience a growth in the frequency and intensity of droughts. This doctoral dissertation consists of two studies exploring possible health issues that might arise from an increase in water scarcity. In the first study, we investigate the effects of abrupt measures implemented by the water company of the Metropolitan Region of São Paulo to cope with an unexpected drought that affected the region during 2014 and 2015. Even though the company manages a robust water and sanitation system, the severity of the drought drove the local water supply services to the brink of a collapse, and a series of strict measures had to be implemented to guarantee the supply to the population. During this drought, an important measure was to reduce the water pressure within the distribution network to avoid losses through leakage points. Our analysis suggests that, although important, this policy had the unexpected consequence of increasing the hospitalization by fecal-oral transmitted diseases. Consistent with the literature, we find evidence that the low-income children and elderly were the most affected groups. Analyzing variations in bottled water prices, we also find evidence that the population engaged with defensive behaviors against the water supply issues, which seems to have influenced the observed health outcomes. In the second study from this dissertation, we discuss another consequence of water scarcity that arises in places that, like Brazil, rely on hydroelectric power plants to supply their energy demand. In these regions, a drought might shift the energy production toward a more intensive use of thermoelectric power. Mounting evidence suggests that this energy source has a negative impact on the health outcomes of the population, and coal-fueled power plants are usually at the center of discussion due to their high emission rates. Still, other types of thermoelectric power plants might also be harmful to the health of the surrounding population. In this work, we investigate if different types of thermoelectric power plants affect the health of the population living nearby. We find evidence that coal-fired power plants are the most harmful for health, increasing the infant mortality rate in nearby communities. Our findings also robustly suggest that female children are more affected than males, but the results are probably underestimated, since we also find evidence of fetal selection.

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

1 Introduction

It is widely accepted by the scientific community that the current path of Greenhouse Gas (GHG) emissions and land use will have a considerable impact on the world's climate conditions. Significant weather variations have already been registered in the last decades and are likely to intensify in the coming years. The Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2013) summarizes different scenarios of climate conditions and extreme weather events such as heatwaves, extreme temperature and precipitation, floods, droughts, and others. It points to a large spatial heterogeneity of these events, with different levels of uncertainty surrounding them. Scientists are fairly confident that the temperature and extreme precipitations are rising and will continue to do so. However, there is less confidence regarding a global trend for other events such as droughts, which will likely affect some regions more than others. The Brazilian northeast and southern Africa are low-income regions that, according to some scenarios, could observe an increase in the intensity of droughts. Under these projected climate scenarios, the affected regions might experience further drawbacks in their pursuit of the United Nations' Sustainable Development Goals (SDG).[1.0.1]

Stanke et al. (2013) summarize the channels through which a drought might affect morbidity and mortality rates. They argue that water shortages, migration, and impacts on livelihood (loss of crops, livestock, fishery, among others) can influence the incidence of illnesses related to malnutrition, mental health, water, vector-borne and airborne diseases. This dissertation investigates two infrastructure bottlenecks that could be severely impacted by the changing climate, influencing societies' health conditions.

After this introduction, the second chapter of this work studies a severe and unexpected drought that affected the Metropolitan Region of São Paulo, the largest urban agglomeration from South America. This extreme weather event drove the region's water and sanitation services to the brink of a collapse. The local water

---

[1.0.1] United Nations (2018), IPCC (2013).

company implemented a series of strict measures to guarantee water supply to the population, but some of these measures had unintended health consequences. In this study, we analyze the health effects of the policies implemented by the water company to cope with the drought. We also explore how the population engaged with defensive mechanisms against the water issues and whether this behavioral factor influenced the observed health outcomes.

The third chapter of this dissertation analyzes a different consequence of a drought that might appear in regions highly dependent on hydroelectric power plants. The water scarcity reduces the hydroelectric power plants' capacity, and thermoelectric facilities usually compensate for this reduction. We investigate if thermoelectric power plants moved by different fuels impact the health outcomes of nearby communities. We analyze the consequences of thermoelectric facilities on early childhood health and infant mortality rates.

The fourth chapter concludes this work by discussing how the main findings contribute to a deeper understanding of the health consequences of climate changes on regions expected to observe an increase in the frequency and intensity of droughts.

2 Health Consequences of Drought-Related Water Supply Policies: Evidence from São Paulo, Brazil

2.1 Introduction

With the ongoing process of climate change, scientists expect that some parts of the world will experience an increase in the frequency and intensity of droughts. As a large share of the world's population and economic activities are located in urban centers, it is essential that these regions adapt to the new climate conditions. Urban agglomerations face many threats related to climate change, such as rising sea levels, heatwaves, flooding, air pollution, and droughts. The increase in frequency and intensity of droughts is particularly unsettling for local governments because they must guarantee proper access to quality water to the population. Modern societies are starting to learn about mitigation and adaptation policies against severe water scarcity, and the consequences of not learning it fast enough are still not well comprehended. Many regions currently rely on a robust water infrastructure with enough capacity to support society's demands even in a situation of water stress. Still, the new pattern of droughts caused by climate change can intensify water scarcity, forcing these regions to adapt to the new climate conditions.[2.1.1]

Because of the large investments needed to adjust supply, water providers have little available short-term options to face a sudden decrease in water availability. This study explores the possible health consequences of policies used to cope with a decrease in water availability in the Metropolitan Region of São Paulo (MRSP). The MRSP is a large urban agglomeration from Brazil that contains 10% of the country's population but produces around 20% of its Gross Domestic Product (GDP).[2.1.2] Although the region has a robust water infrastructure, it experienced a severe drought that led the local water company to implement a series of strict measures to guarantee the supply to the population, and to avoid the complete

---

[2.1.1]IPCC (2013), IPCC (2014).

[2.1.2]Antunes (2018).

depletion of the Cantareira System, the biggest and most important water system of the region.

One important measure to reduce water use during the drought was the pressure reduction in the distribution network. This is an ordinary policy used by water companies to prevent losses through leakage points, but the local water company used it more intensely during the water scarcity period. The water company stressed that this policy was crucial to saving water during the drought (SABESP, 2015b). Despite that, laboratory experiments suggest that pressure changes and the interruption of water supply in the distribution network can compromise the quality of the water delivered to the final user because it might lead to the intrusion of external contaminants through cracks in the pipes.[2.1.3]

Another important water supply policy implemented to expand the capacity of the Cantareira System was the use of the so-called Technical Reserve, a portion of the reservoir below the level of the water collection pipes that had to be pumped up to be used. The Technical Reserve was popularly known as the "dead volume" of the Cantareira System and it had never been used to supply the MRSP. Even without clear evidence of the direct relationship between water from the Technical Reserve and health, salience on the local media led the population to mistrust the newly available water. This mistrust, in turn, may have induced the residents supplied by the Technical Reserve to invest in technologies to improve the quality and availability of water, with possible positive impacts on health.

In this study, we investigate whether the pressure reduction affected hospitalization by fecal-oral transmitted diseases in the MRSP. We also explore the extent to which the Technical Reserve and behavioral factors might have influenced the health consequences of the drought-related water policies.

This work relates to two streams of the literature. The first one explores the linkages between health and access to quality water. Mangyo (2008) and Jalan e Ravallion (2003) use survey information to investigate the health effects of access to water in rural China and India. Their results suggest that improvements in access to quality water improved child health and reduced the prevalence of diarrhea,

---

[2.1.3]Fontanazza et al. (2015).

but these benefits might not be observed in households with uneducated mothers (MANGYO, 2008). The evidence also suggests that child mortality, a more drastic health outcome, reduces after improving access to quantity and quality water.[2.1.4] Zhang (2012) finds evidence that improved water quality reduces the incidence of illness in adults, and it also has a positive effect on the weight-for-height ratio for both adults and children. In addition to direct health benefits, some scholars found evidence of a long-term positive impact of water quality on school performance, especially for female students.[2.1.5]

Ashraf et al. (2017) argue that expanding the infrastructure might not be enough to guarantee all the health benefits from access to adequate water. They investigate the effect of inadequate maintenance of the water infrastructure in Lusaka, Zambia, on health outcomes, economic activity, and time use. Large unanticipated water supply problems can eliminate most of the benefits from the access to water infrastructure because households might not invest in alternative purification methods if they trust the water received is adequate. The results from Ashraf et al. (2017) suggest that water outages are related to adverse health outcomes in children under five years of age (an increase in diarrhea, respiratory infection, and a decrease in vaccination in children under one year of age). They also find an adverse effect of the outages on economic activity and an increase in time girls spend doing their chores.

The second stream of literature related to this work suggests that the population defensively reacts to an increase in pollution, and failing to consider this avoidance behavior can lead to biased estimates of the impact of pollution on health (NEIDELL, 2009). Currie et al. (2013) note that there might exist a selection of households more affected by inadequate water provision. Their work shows that violations in the quality of water might affect birth outcomes, but more educated mothers are more likely to move away from the districts exposed to the violations. This is a particularly serious issue because the literature suggests that exposure to pollution in the early-childhood might affect health and human capital formation

---

[2.1.4]Rocha e Soares (2015), Gamper-Rabindran, Khan e Timmins (2010), Galiani, Gertler e Schargrodsky (2005).
[2.1.5]Zhang e Xu (2016).

throughout the whole life of the exposed individuals (CURRIE et al., 2014). Graff Zivin, Neidell e Schlenker (2011) find more evidence of defensive behavior against low-quality water. They used data from a large grocery chain in California and Nevada to find evidence of an increase in sales of bottled water as a response to violations in the quality of water provided by the public utility company. Avoidance behavior can also have a negative impact on populations' health outcomes. Studying outbreaks of Dengue in Fortaleza, Brazil, Pontes et al. (2000) find that during a drought, an increase in the abundance of the Aedes aegypti mosquito might be associated with households storing water at home, especially when public efforts to suppress the disease vectors are interrupted. Avoidance behavior might be an important factor in the drought of the MRSP. The intense media coverage of the weather shock increased the salience of water-related problems in the region, especially in the neighborhoods supplied by the Cantareira System, which might have induced the population to adapt to the new situation.[2.1.6]

We contribute to this literature by investigating the unintended health consequences of measures adopted under a severe drought to guarantee water supply to the population. Unlike most of the literature, our region of interest has a robust urban infrastructure that seemed adequate to cope with previously experienced droughts. We focus the investigation on the health consequences of the pressure reduction policy. Still, relying on time and spatial heterogeneities of the policies, we also analyze the health impacts on the population supplied by the Cantareira System and discuss how avoidance behavior affected the observed results.

Our results suggest that, although important to save water, the pressure reduction policy had a negative effect on health outcomes. Hospitalization by fecal-oral transmitted diseases increased in regions exposed to this measure, especially in low-income neighborhoods.[2.1.7] The results also suggest that the low-income population supplied by the Cantareira System observed a decrease in hospitalization by fecal-oral transmitted diseases. This seems to be explained by a more active

---

[2.1.6]As discussed in Appendix 2.B.

[2.1.7]Low-income neighborhoods are defined as Census Tracts with more than 50% of the households living with monthly income below or equal to a minimum wage per capita.

engagement with defensive behavior, as we also observed an increase in the price of bottled water in these regions. Although our analyses suggest a positive impact of the pressure reduction policy on hospitalization, we are not able to completely decouple the specific effects of variations in the quality and quantity of water from the effects of people's behavior. We find evidence that these are all plausible explanations, but we can not fully distinguish between these causes.

The remainder of this work is organized as follows. Section 2.2 presents details about the water system that supplies the MRSP and discusses the main efforts from the water company to guarantee water supply to the population. Section 2.3 presents the data used in this analysis, and Section 2.4 discusses the empirical strategy adopted. Section 2.5 shows the main results. Section 2.6 discusses differential effects observed in the Cantareira-supplied regions and the possibilities of population's engagement with avoidance behavior. Finally, Section 2.7 concludes this work.

2.2   Background

The Metropolitan Region of São Paulo is home to around 20 million inhabitants and comprehends 39 municipalities. The local water company, SABESP (*"Companhia de Saneamento Básico do Estado de São Paulo"*), is a mixed capital company with the state's government being the majority shareholder. SABESP supplies water for 35 municipalities from the MRSP and operates in 365 municipalities in the State of São Paulo (SABESP, 2015a; SABESP, 2015b).

The water company manages the Integrated Metropolitan System (SIM) that provided water to the MRSP through eight water reservoirs.[2.2.1] The largest water production facility, called Guaraú Water Treatment Facility (*"ETA Guaraú"*), uses water from the Cantareira System's reservoirs. Before the drought, it supplied water to about half of the population of the MRSP, with a water treatment capacity of up to 33 m^3/s (SABESP, 2015a).

[2.2.1]Alto Cotia, Baixo Cotia, Alto Tietê, Guarapiranga, Rio Claro, Rio Grande, Ribeirão da Estiva, and Cantareira (SABESP, 2015b).

In 2014 and 2015, the MRSP was affected by a severe drought that put stress on the whole water supplying system. This region is marked by a well-defined wet period, from December to February, and a dry period during winter, from June to August. During the wet season of 2013/2014 and 2014/2015, the region experienced a large precipitation deficit, which was aggravated by the fact that previous years were also below the historical average of precipitation for the wet period (COELHO; CARDOSO; FIRPO, 2016). This water scarcity was remarkably intense in the Cantareira System. In the period of 2013/2014, it registered only 53% of the historical average precipitation, reaching the lower amount of rain in more than 80 years (SABESP, 2015a). As shown in Figure 2.2.1, with this adverse weather scenario, the level of water in the reservoirs dropped sharply, and the water company, along with the whole population, had to struggle to adapt to a situation of severe water scarcity.[2.2.2]

Figure 2.2.1 – Level of the water reservoirs in recent periods
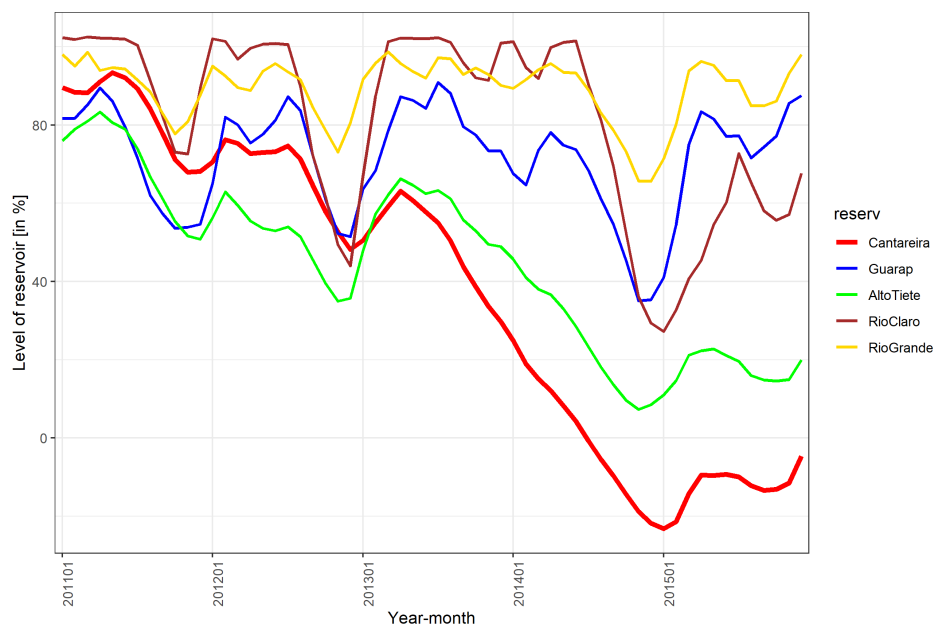
Notes: Figure built by the authors with information from the website of the water company. This figure shows the water availability in relevant reservoirs supplying the MRSP. When the level of the reservoirs from Cantareira System was not available, we interpolated it based on the water availability and its original capacity.

[2.2.2]The negative level of water for the Cantareira System's reservoirs shows the use of the Technical Reserve.

The water company implemented strict actions to guarantee water supply to the population and to prevent the complete depletion of the Cantareira System' reservoirs. Some of these actions were: (i) rearrange the water distribution system so that regions previously supplied by the Cantareira System could also be supplied by other systems; (ii) as of May 2014, the water company started to withdraw water from the Technical Reserve of the Cantareira System;[2.2.3] (iii) the company intensified programs to prevent losses in the system, which included a reduction in the pressure of the distribution network to avoid leakage-related losses; (iv) the company implemented a financial stimulus with bonuses and penalties to encourage clients to save water, but only the latter seems to have been effective, as discussed by Grover e Lucinda (2020); and (v) a wide variety of institutional actions, like advertising to the population about the drought and how to save water.[2.2.4]

2.3 Data

To assess health issues arising in the MRSP during the drought, we used data on hospitalization by water-related diseases taken from the Brazilian Ministry of Health through its system of information (Datasus). This database allows us to identify the zip code of residence of each patient whose hospitalization was paid by the public health system. The data also contains information on the patient's age, gender, days of permanence in the hospital, main disease that led to the hospitalization, among others. This information comes specifically from the Hospital Information System (*"SIH - Sistema de Informações Hospitalares"*) that registers every accepted request made by health facilities to be refunded by the public health system. The main type of disease analyzed throughout this study is the incidence of fecal-oral transmitted diseases with definition based on FUNASA

[2.2.3]In May/2014, when the Technical Reserve 1 was added to the Cantareira System's reservoirs, the total water volume increased 182 million cubic meters (from 982 to 1,164). With Technical Reserve 2, that started to operate in October 2014, another 105 million cubic meters were added to the system.

[2.2.4]While actions 1 through 3 are related to the supply side of the water market, the last two actions can be related to the demand of water.

([2010](#)).[2.3.1]

Socioeconomic characteristics of the MRSP come from the 2010 Brazilian Census of the Brazilian Institute of Geography and Statistics (IBGE), which allows us to work at the Census Tract (CT) level in the regions of interest. To map each zip code from Datasus to a CT of the MRSP, we used georeferenced databases of streets and CTs provided by CEM ("*Centro de Estudos da Metropole*"[2.3.2]). In this mapping, around 10% of the hospitalizations could not be mapped to a CT, so we dropped them from the sample. With information from hospitalization, taken from Datasus, and the population in each CT, taken from the 2010 Census, we are able to build our main output variable, which is the hospitalization rates, given by the number of hospital admissions in a specific month and CT divided by the population of the CT normalized by a hundred thousand. Figure 2.3.1 shows that the hospitalization rates by fecal-oral transmitted diseases increased in the years 2014 and 2015, especially in low-income neighborhoods exposed to the pressure reduction policy.

[2.3.1]Diseases classified by the 10th revision of the International Classification of Diseases by the codes $A00 - A04$, $A06 - A09$, and $B15$.

[2.3.2]http://centrodametropole.fflch.usp.br/en

Figure 2.3.1 – Hospitalization rate by fecal-oral transmitted disease
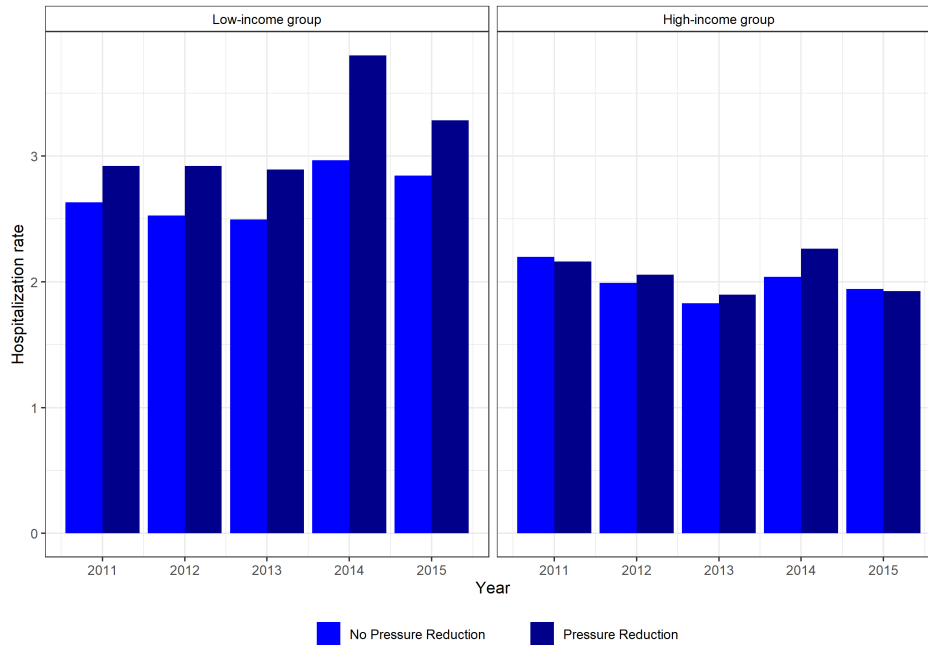
Notes: Built by the authors with information from Datasus, IBGE, and georeferenced maps. This figure shows the hospitalization rates by fecal-oral transmitted diseases in different regions of the MRSP.

To create the main treatment variable, we had to evaluate whether or not the pressure reduction policy was targeting each CT and to find out which CTs received water from the Cantareira System during the drought. Both of these variables were built by georeferencing maps from the water company's publicized reports and regulations.

To evaluate which CTs were exposed to the pressure reduction measure, we georeferenced a map disclosed by the water company showing regions under the influence of Pressure Reduction Valves (PRVs).[2.3.3] This treatment definition contains some measurement error that can not be purged from the variable with the available data. For example, it is possible that regions not influenced by the PRVs also experience a reduction in the pressure of the distribution system, or had PRVs installed after the disclosure of the map. Even regions under the influence of the PRVs might have experienced different pressure dynamics through the drought.

[2.3.3]The water company disclosed this information as a response to deliberation 545/2015 from the regulatory agency.

For instance, we do not know how many hours a week each region experienced the pressure reduction and neither how much was the pressure reduced in each neighborhood (pressure can drop to zero or to half the usual value). Regions under the influence of the PRVs disclosed in the map are considered the most likely exposed to the pressure reduction policy, and, therefore, they will be considered as the treated regions. We also consider that this pressure reduction policy started to operate in February 2014, the same month as the first largely publicized action of the water company to save water - a bonus program aiming at clients living within the region supplied by the Cantareira System (SABESP, 2015a).

Water scarcity was different in each water reservoir supplying the MRSP (as shown in Figure 2.2.1). Therefore, it is also important to consider which region was supplied by each of the water systems. To gather this information, we used a 2015 report (SABESP, 2015a) that presents a map showing areas supplied by five water reservoirs in three different periods (December of 2013, September of 2014, and March of 2015). Georeferencing those maps, we were able to cross information with georeferenced Census Tracts from IBGE and selected only CTs that were supplied by the same water system during the whole analyzed period.

Because the MRSP is a large and heterogeneous region, we had to overcome a couple of challenges to build the database for this study. Initially, some municipalities from the MRSP do not have well-defined zip codes. Streets from these municipalities are all registered under the same few zip codes. In these municipalities, allocating hospitalization to a CT would be unreliable because every hospital admission would be equally allocated among every CT within the city (because they all share the same few zip codes).[2.3.4] Another challenge is that some CTs changed the supplying water systems throughout the drought. In these regions, we do not have precise information about the supplying system in each period of time, so we risk allocating these CTs to the wrong water reservoir.[2.3.5] Besides that, for

---

[2.3.4]Black regions in Figure 2.G.1 show these municipalities with few zip codes.

[2.3.5]We consider that two groups of CTs changed the water system. The first one are those that appeared to be supplied by different reservoir in each of the periods from the maps (maps from December of 2013, September of 2014, and March of 2015) of the 2015 report from the water company (SABESP, 2015a). The second group contains CTs from the *Paulista Avenue* and the *Pinheiros* regions. They were originally supplied by the Cantareira System, but the report mentioned above (SABESP, 2015a) indicated that they would start to be supplied by other water

some CTs we do not have sufficient information about the water system.[2.3.6] To avoid considering the wrong assumptions regarding the water condition in each neighborhood, we excluded from the main sample all these CTs with incomplete and unreliable information about the supplying water system.

Figure 2.3.2 presents the details of the main sample used throughout this study after all the exclusions discussed in the previous paragraphs. Dark blue and dark green areas in the north of the MRSP contain CTs constantly supplied by the Cantareira System, while those in light colors received water from some other water system. CTs painted in blue (dark or light) are those exposed to the pressure reduction policy. White regions contain CTs that changed the water system throughout the drought (the central region) or CTs in which we have no precise information about the supplying system. Regions in black are those in municipalities with few zip codes that were excluded from the sample.

---

systems throughout 2015.

[2.3.6]We consider that we do not have precise information for two groups of CTs: CTs not allocated to any supplying system in the report from 2015 (SABESP, 2015a), and CTs allocated to the Cantareira System in this 2015 report, but not in other reports from the company. A 2011 environmental impact report of the São Lourenço Production System (*"Relatório de Impacto ao Meio Ambiente do Sistema Produtor São Lourenço"*) had a map showing regions supplied by each water reservoir (SABESP, 2011). We georeferenced regions from the Cantareira System in this report to further check for regions allocated to the wrong water system.

Figure 2.3.2 – Sample selection
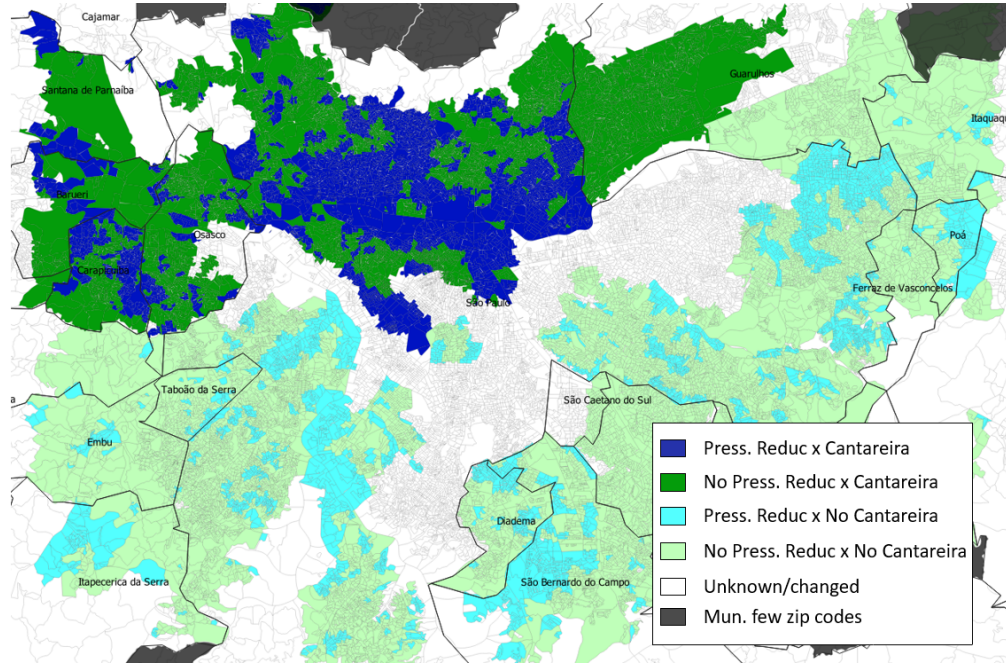
Notes: This figure presents the main sample selection used throughout the analysis.

Appendix Section 2.G presents different maps of the MRSP and the selected sample,[2.3.7] and the Table 2.A.1 compares characteristics of the CTs included in the main sample with those excluded from the analysis. The region used in this study is larger than that not used, both in terms of population and number of Census Tracts (but not in land cover).

Table 2.3.1 compares characteristics of the CTs from the main sample. Column 1 shows the characteristics of CTs that receive water from the Cantareira water system but that are not exposed to the pressure reduction policy. Column 2 shows those regions receiving water from the Cantareira that are exposed to the pressure policy. Column 3 presents the characteristics of CTs that do not receive water from the Cantareira System and that are not exposed to the pressure

---

[2.3.7]Figure 2.G.1 shows all the 39 municipalities from the MRSP, highlighting the cities excluded from the analysis and the regions most likely exposed to the pressure reduction policy. Figure 2.G.2 shows the full MRSP highlighting in red the CTs included in the main sample. Figure 2.G.3 shows which regions were considered to be supplied by which water reservoir throughout the whole drought, and Figure 2.G.4 shows the average altitude of each CT from the sample.

reduction policy. Finally, column 4 shows the regions that do not receive water from the Cantareira System but that are exposed to the pressure reduction.

Initially, we note that the regions presented in column 3 (outside Cantareira System and not exposed to pressure reduction) are much larger in terms of inhabitants and number of Census Tracts. The regions outside the Cantareira System's influence also concentrate a larger share of the low-income population, both outside (column 3) and inside (column 4) the pressure reduction area. The population living in regions receiving water from the Cantareira System (columns 1 and 2) are slightly older, with higher income, more educated, and better served by urban utilities than the rest of the sample. The access to the public water system and to garbage collection is close to 100% in every region from the sample (not shown in the table), but access to the sewage system is slightly higher in the neighborhoods receiving water from the Cantareira System.[2.3.8] This shows that the water infrastructure in the MRSP is already robust compared to those investigated in most of the papers from the literature related to this study. Finally, CTs exposed to the pressure reduction policy (columns 2 and 4) have a slightly lower altitude.

[2.3.8]Being connected to the sewage system is different from having the sewage treated. Although the share of people connected to the sewage collecting system is high, the share of treatment waste water is low.

Table 2.3.1 – Descriptive statistics

| | Cantareira | | No Cantareira | |
| | No Pr. Reduc | Pr. Reduc | No Pr. Reduc | Pr. Reduc |
Variables	(1)	(2)	(3)	(4)
Low-income CTs	1429	1082	5223	1987
High-income CTs	2456	2179	4506	1466
Low-income pop [MI]	1.09	0.79	3.76	1.36
High-income pop [MI]	1.5	1.3	2.74	0.86
Total fecal-oral hosp. [K]	3.41	3.61	9.92	3.64
Shr pop 0-5 years	0.06	0.06	0.07	0.07
	(0.02)	(0.02)	(0.02)	(0.02)
Shr pop above 60	0.11	0.12	0.09	0.09
	(0.06)	(0.06)	(0.06)	(0.05)
Income pc [2018 K BRL]	1.67	1.79	1.28	1.29
	(1.5)	(1.82)	(1.34)	(1.51)
Literacy rate	0.96	0.96	0.95	0.95
	(0.04)	(0.04)	(0.04)	(0.04)
Shr HH sewage	0.89	0.91	0.88	0.89
	(0.23)	(0.2)	(0.24)	(0.24)
Average height [m]	777.35	763.13	790.22	774.76
	(29.72)	(25.65)	(24.16)	(26.67)

Notes: This table shows the main characteristics of the preferred sample dividing it between the regions analyzed. Column 1 shows CTs supplied by the Cantareira System but not exposed to the pressure reduction policy. Column 2 presents the characteristics of CTs that receive water from the Canteira System and are exposed to the pressure reduction. Column 3 shows CTs not supplied by the Cantareira System and not exposed to the pressure reduction. Column 4 presents the characteristics of the CTs not supplied by the Cantareira System but exposed to the pressure reduction policy. The data comes from the 2010 IBGE Census and from the INPE's Topodata project. The income per capita is given in thousands of 2018 BRL. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita.

To gather information on the level of water in each of the main water reservoirs from the MRSP, we averaged daily information taken from the water company's website. We used data from the Topodata project, of the National Institute for Spatial Research (INPE), to build the average altitude of each Census Tract. This data contains information about altitude in a grid with a resolution of around 90 meters, and the average altitude is calculated based on the average grid that falls within each CT, or the closest grid point. Finally, the period of the analysis is from January 2011, three years before the beginning of the drought, until the end of 2015, when the Technical Reserve from the Cantareira System

stopped being used (two years after the beginning of the drought).

2.4   Specification

In this study, we are interested in evaluating if the pressure reduction policy implemented by the water company in the MRSP affected health outcomes of the population. We rely on the heterogeneous spatial distribution of the water-saving measure to investigate its impact on hospitalization by fecal-oral transmitted diseases. Our first specification presented in Equation 2.4.1 is a difference-in-differences regression.

$$y_{it} = \beta_1 d_t LowPress_i + \theta L_{it} + \gamma_{mt} + \delta_i + \epsilon_{it} \qquad (2.4.1)$$

Where y_{it} is the hospitalization rate on CT i at month t. $LowPress_i$ and d_t are indicator variables equal to 1, respectively, if CT i is exposed to the pressure reduction policy and if at month t the policy had already started. L_{it} controls for the monthly average level of the water reservoir supplying CT i at period t, to account for diseases caused not by the policy under scrutiny but by the low level of the water in the supplying reservoir. γ_{mt} is a municipality by period fixed effect that captures actions implemented in each period t by the municipality m to which CT i belongs. δ_i is a Census Tract fixed effect to capture time-invariant characteristics of that CT. ϵ_{it} is an idiosyncratic error. The coefficient of interest is β_1 that captures the relative increase in hospitalization by a hundred thousand inhabitants from patients exposed to the pressure reduction policy.

As discussed in the data section, our pressure reduction variable includes measurement errors that could not be purged away with the available information, and we might be mistakenly allocating some CTs to the treatment or control status. Therefore, our estimations might be suffering from attenuation bias. Besides that, the estimated coefficients capture not only the effect of the quality of the water on health outcomes, but the net effect of it bundled with any type of avoidance behavior that the population might have engaged with. The underlying assumption of the

difference-in-differences method is that health outcome trends would have been the same in control and treatment groups, had the policies not being implemented. To investigate this hypothesis, in Section 2.5.3 we conduct robustness checks using hospitalization rates by other types of diseases unrelated to water.

We estimate the regressions with a two-way cluster robust covariance matrix at the month and Census Tract levels to account for: (i) within Census Tract serial correlation over time; and (ii) spatial correlation between the Census Tracts in a given month. We also weigh the regressions by the number of inhabitants from each Census Tract.

To better understand the dynamic effects of the policy under investigation, the difference-in-differences analysis will also be replicated with an event-study framework interacting the treatment with dummies for lags and leads from the month considered as the beginning of our treatment (January of 2014). This approach intends to capture policy heterogeneous responses along the time dimension. Time heterogeneity may arise, for instance, if people realize after a few months that the quality of water is getting worse and react by taking preventive measures. This could result in an initially high impact of the policy on health outcomes, followed by a decrease in treatment effect. On the other hand, it is also possible that the length of exposure to low-quality water has cumulative effects on the population's health, so the usage of the public health system might be higher after some months of the policy in place. The effect of the media coverage of the water issues is also relevant to time heterogeneity since it can influence peoples' behavior throughout the drought.

Analyzing treatment effect before the policy implementation in an event-study setup is also akin to a parallel trend test, therefore, with this specification, we further analyze the validity of the difference-in-differences hypothesis. The estimated equation for the event-study is given by Equation 2.4.2:

$$y_{it} = \sum_{\tau \in (pbp, pap)} (\varphi_\tau D_{\tau t} LowPress_i) + \theta L_{it} + \gamma_{mt} + \delta_i + \epsilon_{it} \qquad (2.4.2)$$

Where y_{it} stands for the same health outcomes as in the difference-in-

differences analysis. *pbp* and *pap* stand for periods before and after the policy was implemented (that varies from -36 in January of 2011 to +23 in December 2015). $LowPress_i$ equals 1 if CT i is exposed to the pressure reduction policy. $D_{\tau t}$ is an indicator variable that equals 1 if period t is τ months away from the beginning of the policy. Before the policy is implemented, τ is negative, it is equal to zero in January of 2014, and it is positive thereafter. γ_{mt}, δ_i, and L_{it} are fixed effects and a control like in the previous specification. ϵ_{it} is an idiosyncratic error. The coefficient of interest is φ_τ, which measures the relative variation of hospitalization by a hundred thousand inhabitants τ months away from the policy implementation for patients living in CTs exposed to the pressure reduction policy.

2.5   Results

2.5.1   Main result

Table 2.5.1 presents the main results of this analysis. Panel A and B, show the results for, respectively, low and high-income populations. Columns 1 through 3 use the rate of hospitalization as the dependent variables, while in column 4, the dependent variable is in level (hospitalization by month in each CT).

Table 2.5.1 – Hospitalization by fecal-oral transmitted disease

	Dependent variable			
	Rate			Level
	(1)	(2)	(3)	(4)
Panel A: low-income neighborhoods				
lowPress:dt	0.282**	0.417***	0.361**	0.003**
	(0.121)	(0.133)	(0.149)	(0.001)
Mean dep. var	2.83	3.03	3.03	0.03
Observations	805,740	583,260	583,260	583,260
Adjusted R^2	0.0004	0.001	0.033	0.045
Panel B: high-income neighborhood				
lowPress:dt	0.112	0.235**	0.192*	0.002**
	(0.103)	(0.099)	(0.104)	(0.001)
Mean dep. var	2.02	2.05	2.05	0.02
Observations	934,320	636,420	636,420	636,420
Adjusted R^2	0.00000	0.00004	0.104	0.100
Preferred sample		x	x	x
CT FE			x	x
Period x Mun FE			x	x
Level control			x	x

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Columns 1 through 3 use the rate of hospitalization by a hundred thousand inhabitants as the dependent variable, while column 4 uses hospitalization in level. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita.

Column 1 is a simple difference-in-differences without any controls or fixed effects. Because in this model we are not controlling for the level of the reservoir supplying each CT, we can use the whole MRSP without the need to exclude those CTs that either changed the supplying system or for which we have no information. This larger sample is presented in Figure 2.G.1. In this specification,

we only exclude from the sample CTs from municipalities with few zip codes. In column 2, we repeat the specification without controls of fixed effects but using only the preferred sample presented in Figure 2.3.2. Columns 3 and 4 present our benchmark specification, with the preferred sample and all the controls and fixed effects discussed in the previous sections. While in column 3, the dependent variable is the rate of hospitalization, in column 4, we use the variable in levels.

The effect of the pressure reduction policy in hospitalization by fecal-oral transmitted disease is positive in every specification, and it is always stronger in low-income neighborhoods.[2.5.1] When we move from the specification with the whole metropolitan region to that with our preferred sample (from column 1 to 2), the effect increases in magnitude and becomes more significant. Results from the benchmark specification (column 3) suggest that the effect of the pressure reduction policy in the rate of hospitalization by fecal-oral transmitted diseases accounts for around 10% of the average hospitalization rate, with low-income neighborhoods observing slightly larger effects.

Figure 2.5.1 presents the estimated coefficients from the event-study aggregating the months at the trimester level. We notice that the average coefficients, represented by the dashed horizontal red lines, are larger after the beginning of the pressure reduction policy than before it, especially for low-income neighborhoods. In these regions, we observe an increase in hospitalization in the four trimesters after the beginning of the policy, then the coefficient drops to a value close to zero again. The event-study, presented in appendix Figure 2.E.1, uses the specification with monthly periods instead of trimesters. It shows a similar pattern but with noisier coefficients. The increase in the coefficients after the policy intervention followed by an eventual drop is consistent with the population increasingly engaging with defensive behavior.

[2.5.1]Results are similar when using a larger spatial aggregation. In Appendix 2.F, we repeat this analysis using Weighting Areas (an aggregation of Census Tracts) as the unit of observation. The results are similar, but slightly less significant, which might be due to the reduced number of observations.

Figure 2.5.1 – Event-study for the hospitalization by fecal-oral diseases
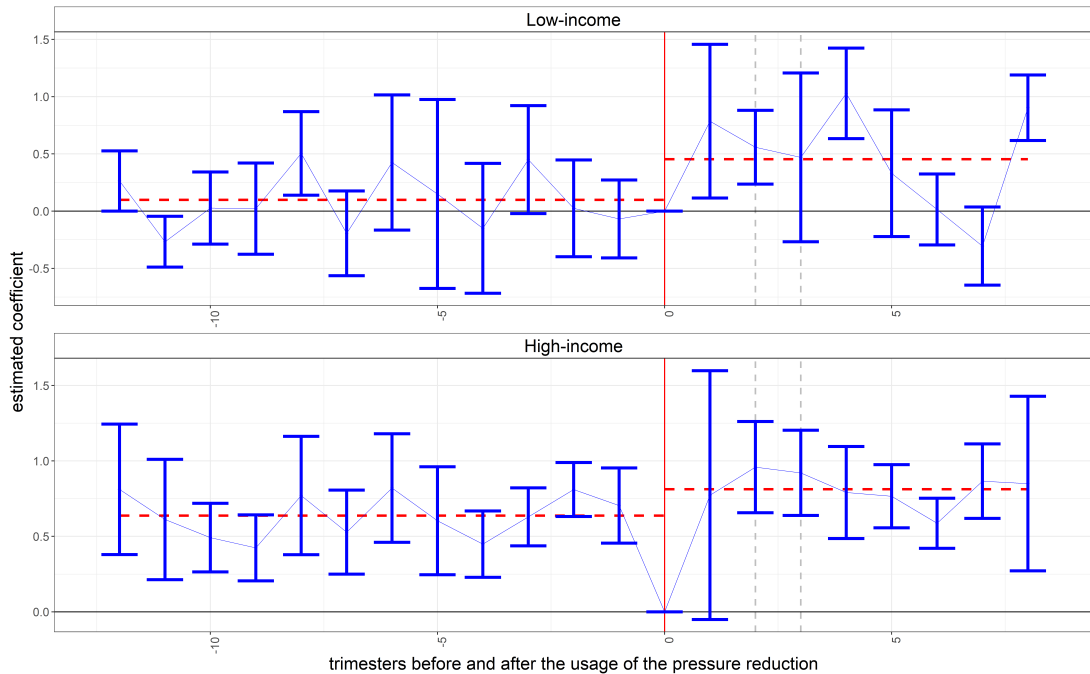
Notes: This figure presents the estimated coefficients for the event-study showing the impact of the pressure reduction policy on hospitalization in each trimester before and after the beginning of the policy.

Before the beginning of the pressure reduction policy, the coefficients seem to be stably varying around a mean. This is an evidence in favor of the parallel trend hypothesis for the treatment and control groups before the implementation of the policy, as discussed in the specification section.

2.5.2 Age heterogeneity

Using our benchmark specification, in Table 2.5.2 we investigate the effect of the pressure reduction policy for different age groups. The dependent variables are defined as hospitalization of people in each age group by a hundred thousand inhabitants from the same age group in each CT. Results suggest that the population under 5 and above 60 years of age are the most affected groups, and the effect is much stronger in low-income neighborhoods.

Table 2.5.2 – Hospitalization by fecal-oral transmitted disease by age group

	Dependent variable			
	Age 0-5	Age 5-18	Age 19-60	Age 60-more
	(1)	(2)	(3)	(4)
Panel A: low-income neighborhoods				
lowPress:dt	2.529**	0.148	0.063	1.190*
	(1.204)	(0.303)	(0.078)	(0.687)
Mean dep. var	18.05	3.18	0.78	5.35
Observations	583,260	583,260	583,260	583,260
Adjusted R^2	0.020	0.012	0.007	0.004
Panel B: high-income neighborhood				
lowPress:dt	0.604	0.330	0.067	0.387
	(1.217)	(0.271)	(0.061)	(0.289)
Mean dep. var	13.56	2.77	0.61	3.29
Observations	628,500	628,500	628,500	628,500
Adjusted R^2	0.088	0.056	0.016	0.011

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each age group. 143 high-income Census Tracts did not contain age information, so they were excluded from the sample. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita. Each column presents a different age group.

2.5.3 Robustness check

Table 2.5.3 shows the main robustness check from this exercise. We use the benchmark specification to investigate if the beginning of the pressure reduction policy affected other diseases that are not directly related to water.

Table 2.5.3 – Hospitalization by other diseases

	Dependent variable				
	Vector	Flu	Neoplasm	Appendicitis	Head injury
	(1)	(2)	(3)	(4)	(5)
Panel A: low-income neighborhoods					
lowPress:dt	0.135	−0.007	−0.056	0.025	0.198
	(0.133)	(0.443)	(0.068)	(0.160)	(0.193)
Mean dep. var	0.57	25.41	0.68	4.37	6.83
Observations	583,260	583,260	583,260	583,260	583,260
Adjusted R^2	0.045	0.154	0.017	0.030	0.050
Panel B: high-income neighborhood					
lowPress:dt	0.129	0.444	−0.150*	0.253*	−0.082
	(0.103)	(0.482)	(0.080)	(0.130)	(0.222)
Mean dep. var	0.45	18.7	0.85	3.14	5.89
Observations	636,420	636,420	636,420	636,420	636,420
Adjusted R^2	0.041	0.337	0.020	0.083	0.245

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita. The dependent variable from each column is the hospitalization rate by a specific disease considering all age groups.

Most coefficients are statistically indistinguishable from zero, and we observe mixed signs in both income groups. The only borderline statistically significant coefficients are for neoplasm and appendicitis in high-income neighborhoods.

2.6 Cantareira System and the avoidance behavior

We previously discussed that the Cantareira System's reservoirs were the most affected by the drought, and that, as of May 2014, the water company started using the Technical Reserve to increase the water capacity of it. In this section, we test whether the population receiving water from this water reservoir experienced different dynamics in hospitalizations by fecal-oral transmitted disease. The first specification, presented in Equation 2.6.1, tests whether the water from the Technical Reserve also affected hospitalization by fecal-oral transmitted diseases.

$$y_{it} = \beta_1 d_t LowPress_i + \beta_2 dcant_t Cant_i + \theta L_{it} + \gamma_{mt} + \delta_i + \epsilon_{it} \qquad (2.6.1)$$

In this specification, $Cant_i$ is an indicator variable that equals 1 if Census Tracts i receives water from the Cantareira System throughout the whole period of the drought and the variable $dcant_t$ equals 1 if period t is after the implementation of the Technical Reserve (equals 1 after May of 2014). Table 2.6.1 shows the results for this specification in columns (1) for low-income and (3) for high-income neighborhoods.

Because the Cantareira System was more affected by the drought, it is also possible that the pressure reduction policy was more intensely used in regions supplied by this reservoir. We test this hypothesis with a triple-differences specification presented in Equation 2.6.2. The interpretation of this model is slightly different from the regular triple-differences used in the literature. Usually, this method is used to control for two sets of unobservable time-varying characteristics, so each control group would be used to purge the variations of a specific characteristic. Here, the triple-differences is used to investigate whether the pressure reduction policy had a different effect inside and outside of the region supplied by the Cantareira System. The model specification is as follows:

$$\begin{aligned} y_{it} =& \beta_1 d_t LowPress_i + \beta_2 d_t Cant_i + \beta_3 d_t LowPress_i Cant_i \\ & + \theta L_{it} + \gamma_{mt} + \delta_i + \epsilon_{it} \end{aligned} \qquad (2.6.2)$$

This specification is similar to the regular differences-in-differences from Equation 2.4.1, but we add another interaction term $Cant_i$ (equals 1 if CT i is supplied by the Cantareira System). The coefficients of interest are β_1, β_2, and β_3. In this triple-differences specification, β_1 is the relative increase in the hospitalization rate from patients exposed to the pressure reduction policy. Similarly, β_2 in the relative increase of hospitalization from patients living within the Cantareira region. Finally, β_3 is the coefficient of the triple interaction, and it measures the relative increase in hospitalization of patients affected by the pressure reduction within the Cantareira region. In other words, this coefficient is the excess impact of being exposed to the pressure reduction within the Cantareira region relative to the impact of it outside this region. Table 2.6.1 shows results for this specification in columns (2) for low-income and (4) for high-income neighborhoods.

Table 2.6.1 – The effect of the Cantareira System

	Dependent variable			
	Rate of hospitalization: fecal-oral			
	(1)	(2)	(3)	(4)
	Low-income		High-income	
lowPress:dt	0.380**	0.439***	0.192*	0.489***
	(0.150)	(0.154)	(0.106)	(0.156)
cant:dcant	−1.095***		−0.010	
	(0.315)		(0.199)	
dt:cant		−0.696*		0.223
		(0.359)		(0.240)
lowPress:dt:cant		−0.190		−0.593**
		(0.271)		(0.242)
Mean dep. var	3.03		2.05	
Observations	583,260	583,260	636,420	636,420
Adjusted R^2	0.033	0.033	0.104	0.104

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. The dependent variable is the rate of hospitalization by a hundred thousand inhabitants for fecal-oral transmitted diseases. Columns 1 and 2 analyze the effects in low-income regions, while columns 3 and 4 are for high-income CTs. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita. Columns 1 and 3 use the specification from Equation 2.6.1, and columns 2 and 4 use specification from Equation 2.6.2.

The results from both specifications suggest that the low-income CTs receiving water from the Cantareira System improved after the beginning of the policy when compared to the rest of the population of the MRSP. Hospitalization by fecal-oral transmitted diseases decreased in low-income neighborhoods after the beginning of the usage of the Technical Reserve (column 1), and the triple-

differences specification presents no evidence that the pressure reduction was more harmful in the Cantareira region. In fact, for high-income neighborhoods, the beginning of the pressure reduction policy actually reduced the hospitalization in regions supplied by the Cantareira System, relative to the effect observed outside this reservoir.

These reductions in hospitalization observed in regions supplied by the Cantareira System might be explained by the population engaging in defensive behavior against the consequences of water issues. Salience about the problems related to the water supply in the Cantareira System was high in the MRSP during the drought, and it is possible that the population reacted to this information by avoiding consuming the water supplied by the water company.[2.6.1] To test this hypothesis, we follow Graff Zivin, Neidell e Schlenker (2011) and test if the policies affected the consumption of bottled water. We have information on the average price of bottled water in different zip codes from the MRSP.[2.6.2] This data was gathered by the Foundation Institute for Economic Research (FIPE) to calculate the region's price index. To allocate bottled water prices from zip codes to CTs, we calculated the average price observed in every zip code within 2 kilometers of each CT, and weighted by the inverse of the distance between the CT and the zip code.

[2.6.1]We discuss the salience and media coverage of the problems related to the Cantareira System in Appendix 2.B.

[2.6.2]We don't observe the volume of water sold in each zip code, but we assume that the price and volume are related.

Table 2.6.2 – Observed effect in bottled water prices

	Dependent variable			
	Within 2 km price of bottled water			
	(1)	(2)	(3)	(4)
	Low-income		High-income	
lowPress:dt	0.022	0.017	−0.035***	−0.028***
	(0.016)	(0.018)	(0.010)	(0.010)
cant:dcant	0.189***		0.036	
	(0.062)		(0.02355)	
dt:cant		0.220***		0.067**
		(0.070)		(0.027)
lowPress:dt:cant		0.003		−0.017
		(0.027)		(0.017)
Mean dep. var.	1.93		1.89	
Observations	203,629	203,629	331,075	331,075
Adjusted R²	0.523	0.526	0.553	0.554

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. The dependent variable is the average price of bottled water within 2 km of each CT. Columns 1 and 2 analyze the effects in low-income regions, while columns 3 and 4 are for high-income CTs. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita. Columns 1 and 3 use the specification from Equation 2.6.1, and columns 2 and 4 use specification from Equation 2.6.2.

Results presented in Table 2.6.2 suggest that bottled water prices increased in low-income regions supplied by the Cantareira System. This result, combined with results in columns 1 and 2 from Table 2.6.1, corroborates with the avoidance behavior hypothesis. For the high-income neighborhoods, the low-pressure region seems to induce a reduction in the prices, but the magnitude of the coefficient

is much smaller than the increase observed in low-income Cantareira regions. Furthermore, it is possible that high-income households invested in more costly defensive actions such as better water filtration and storage systems.[2.6.3] These investments, in turn, might have led to a relative drop in the demand for bottled water. Results are similar when using prices from within 10 km of each CT, as shown in Appendix Table 2.D.2.

These results are suggestive that the population reacted to the intense media coverage about water scarcity in the Cantareira System's water reservoirs by engaging in avoidance behavior. This defensive behavior might have led the more at risk population from the low-income regions to observe a relative reduction in the hospitalization by fecal-oral diseases.

2.7 Discussion

Climate change is expected to increase water scarcity in different parts of the planet. This adverse weather shock can have severe effects on urban agglomerations, posing a risk to water supply services of local governments around the world. In this work, we investigate health consequences of a policy implemented to cope with a severe and unexpected drought that affected a large metropolitan region from a developing country.

In 2014 and 2015, the Metropolitan Region of São Paulo (MRSP) suffered one of the most severe droughts of its recent history. This adverse weather shock drove the water services, which seemed adequate to deal with past weather events, to the brink of a collapse. To manage this situation, the local water company had to implement a series of strict measures to guarantee water supply to the population and to avoid the complete depletion of the Cantareira System, the largest and most important water reservoir from the region. In this work, we analyzed the unintended health consequences of the pressure reduction policy, which consisted of reducing the pressure of water within the distribution network to prevent losses

---

[2.6.3]Figure 2.B.2, from Section 2.B, shows an increase in the volume of internet searches for terms related to defensive mechanisms against droughts. Although not conclusive, this could indicate higher interest in investing in defensive methods.

through leakage points. We also discussed the health dynamics in regions supplied by the Cantareira System and the influence of avoidance behavior.

The main results suggest that the pressure reduction policy increased hospitalization by fecal-oral transmitted diseases. We also found evidence that the population engaged with defensive behavior against the drought, especially in regions supplied by the Cantareira System. The defensive measures seems to have influenced the observed hospitalization rates. In accordance with most of the literature, our results indicate that children under 5 years of age and the elderly population above 60 years old, living in low-income neighborhoods, are the most affected by the water policy. In our benchmark specification, the pressure reduction policy in low and high-income neighborhoods represents, respectively, around 12% and 9% of the hospitalization rates observed in each region.

This study highlights that the sudden implementation of water-supply policies to cope with an unexpected drought can affect the population's health. While the pressure reduction and the Technical Reserve of the Cantareira System can both be considered a water-supply policy, we only found evidence that the former had a negative effect on health. Nevertheless, our evidence on the population's engagement with defensive behavior, along with Grover e Lucinda (2020) argument that consumers react to financial stimulus, suggests that water companies could also rely on demand-side policies to manage a situation of water stress.

It is important to note that, in this study, we only investigate the aggregated impact of the pressure reduction policy, without delving into the specific mechanisms that led to the increase in hospitalization. The quality of the tap water, water shortages, as well as behavioral factors, are all potential channels influencing the observed hospitalization rates, and they all seem to play a role in this negative consequence of the water-saving policy. Further investigations are needed to disentangle these effects and they would contribute to a better understanding of the impacts of droughts in urban agglomerations.

Appendix

2.A Descriptive statistics

Table 2.A.1 – Preferred sample compared to the rest of the MRSP

	Variable	Preferred region	Rest of the MRSP
1	Total population [MI]	13.49	6.17
2	Quantity of Census Tracts	20452	9423
3	Pop. under 1 year of age [thousand]	179.83	80.92
4	Pop. between 1 and 5 years old [thousand]	721.4	324.53
5	Pop. between 6 and 18 years old [thousand]	2931.02	1318.8
6	Pop. between 19 and 60 years old [thousand]	8364.48	3799.5
7	Pop. above 60 years old [thousand]	1288.15	636.7
8	Shr. of pop. under 1 year of age	0.01	0.01
9		(0.01)	(0.01)
10	Shr. of pop. between 1 and 5 years old	0.05	0.05
11		(0.02)	(0.02)
12	Shr. of pop. between 5 and 18 years old	0.22	0.21
13		(0.06)	(0.06)
14	Shr. of pop. between 18 and 60 years old	0.62	0.62
15		(0.04)	(0.05)
16	Shr. of pop. above 60 years old	0.1	0.1
17		(0.06)	(0.06)
18	Inhabitants per CT	658.22	651.52
19		(332.58)	(353.13)
20	CT altitude	780.86	780.51
21		(27.85)	(41.15)
22	Monthly income per capita	1434.01	1641.16
23		(1494.51)	(1756.25)
24	Shr. of dwellings under 1 min. wage	0.49	0.47
25		(0.23)	(0.24)
26	Shr. of dwellings from 1 to 2 min. wage	0.27	0.26
27		(0.09)	(0.1)
28	Shr. of dwellings from 2 to 3 min. wage	0.09	0.09
29		(0.06)	(0.06)
30	Shr. of dwellings from 3 to 5 min. wage	0.07	0.08
31		(0.07)	(0.08)
32	Shr. of dwellings from 5 to 10 min. wage	0.05	0.07
33		(0.09)	(0.1)
34	Shr. of dwellings above 10 min. wage	0.02	0.03
35		(0.07)	(0.08)
36	Literacy rate	0.96	0.96
37		(0.04)	(0.04)
38	HH with water	0.99	0.95
39		(0.06)	(0.18)
40	HH with sewage	0.9	0.82
41		(0.22)	(0.32)
42	HH with garbage	1	0.99
43		(0.03)	(0.04)

Notes: Comparing the characteristics of the main sample with the Census Tracts excluded from it. Standard deviations are presented in parenthesis.

2.B Salience of water issues in the Metropolitan Region of São Paulo

In this section, we discuss the salience of the drought in the media during the period of water scarcity. Figure 2.B.1, shows the number of articles found on the website of *Folha de São Paulo*, a large Brazilian newspaper, for three different categories of search parameters. The first category, Cantareira, contains phrases associated with the Cantareira System. The second category contains generic searches about the drought. The third specifically looks for references to the pressure reduction in the water distribution system.[2.B.1]

This figure reveals some caveats about the information received by the population that could potentially influence awareness of the drought, and, consequently, the observed hospitalization by fecal-oral transmitted diseases. Initially, we notice almost no discussion about the drought or the Cantareira System prior to 2014. However, at the beginning of that year, the water company implemented the first bonus program in regions supplied by the Cantareira System, and, around the same time, news about the drought and about the Cantareira started to appear in the media. News related specifically to the pressure reduction in the distribution system appeared only at the beginning of 2015.

This salience about problems in the Cantareira System might have led the population to engage more intensely with avoidance behavior, explaining the increase in the price of bottled water in this region, as well as the relative in reduction hospitalization by fecal-oral transmitted disease.

[2.B.1]A single article can belong to more than one category and we rely entirely on the search engine from the newspaper to select articles according to input words.

Figure 2.B.1 – Count of articles related to the water issues of the MRSP
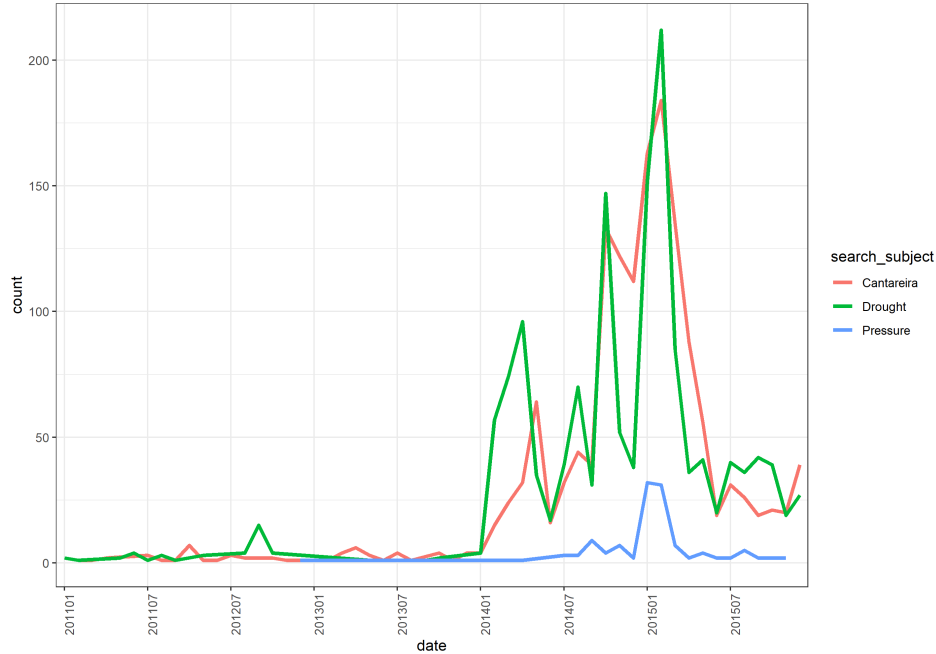
We further discuss population's engagement with defensive behavior by analyzing the dynamics of internet search for certain terms related or not to the water shortage. We use data gathered by Google Trends and investigate internet searches within the State of São Paulo. Google Trends initially provides a search index in a weekly basis. In order to compare the results with other analyses from this study, we aggregate this data to a monthly search index. We analyze this search index for three terms associated to protection against water shortage, and four terms associated to public policy, but not to the drought. The search terms were all in Portuguese, and their translations are: water tank, cistern, artesian well, public transportation, corruption, violence, and traffic.

Based on Bai e Perron (2003), we investigate structural breaks in the time

series of the search index for each term.[2.B.2] Considering a time series with m structural breaks and $m+1$ segments of data (j), we analyze the model given by:

$$search_t = \delta_j + \beta_j x_t + u_t \qquad (2.B.1)$$

where $j = 1, ..., m+1$ and $t = t_{j-1}+1, ..., t_j$. $search_t$ is the search index for a given term within the State of São Paulo in month t. δ_j is the average search index in the segment j and is our variable of interest. x_t controls for search indexes for the same term as $search_t$, but for all states outside the southeast region. This control accounts for changes in the interest for each term that might not be correlated to the drought in the MRSP. We excluded states from southeast because they might have experienced similar patterns of water scarcity. To aggregate the search index of every state, we transformed it to standard deviations and averaged it for each month. For each possible segment of the data, the algorithm calculates the least-square estimates for each coefficient and computes the Residual Sum of Squares (RSS). A dynamic programming algorithm is used to compute the breakpoints that minimize the sum of all RSS.[2.B.3]

The Figures 2.B.2 and 2.B.3 show the estimated structural breaks and corresponding confidence intervals for search indexes of terms related to the drought (Figure 2.B.2) and unrelated to water shortages (Figure 2.B.3). The gray vertical lines indicate the estimated structural break and the red dates on the x axis show the dates associated with the breaks. Red horizontal lines show the confidence interval for each break point calculated by the algorithm. The search indexes for all terms related to protection against droughts have two structural breaks, one in the beginning of 2014 and another in the beginning of 2015. The terms unrelated to droughts either have no structural breaks or the break is before the beginning of the drought.

[2.B.2] Estimation of the structural breaks are implemented by the R package *strucchange* (ZEILEIS et al., 2003).

[2.B.3] For more details, refer to Bai e Perron (2003) and Zeileis et al. (2003).

Figure 2.B.2 – Monthly searches in Google Trend São Paulo State: drought-related terms
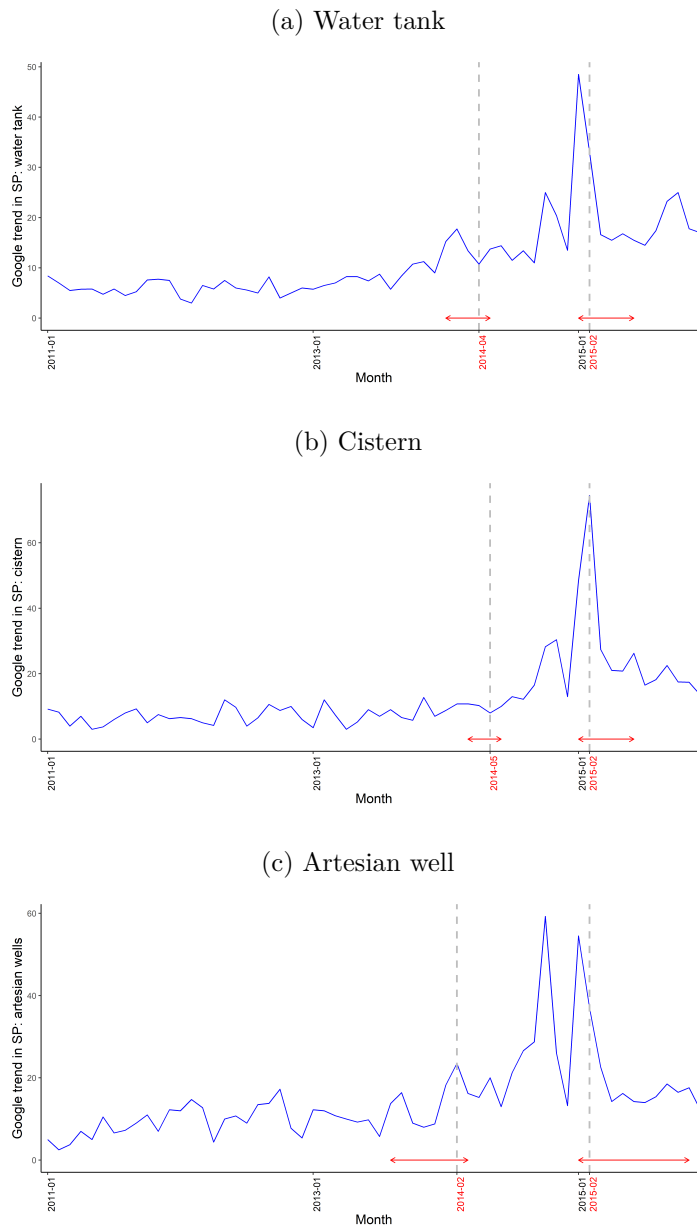
(a) Water tank

(b) Cistern

(c) Artesian well

Notes: These figures show the dynamic in search entries from Google Trends for each specific term related to the drought. The gray vertical lines indicate the estimated structural break and the red dates on the x axis show the estimated date of the break. Red horizontal lines show the confidence interval for each break point calculated by the algorithm.

Figure 2.B.3 – Monthly searches in Google Trend São Paulo State: terms not related to the drought
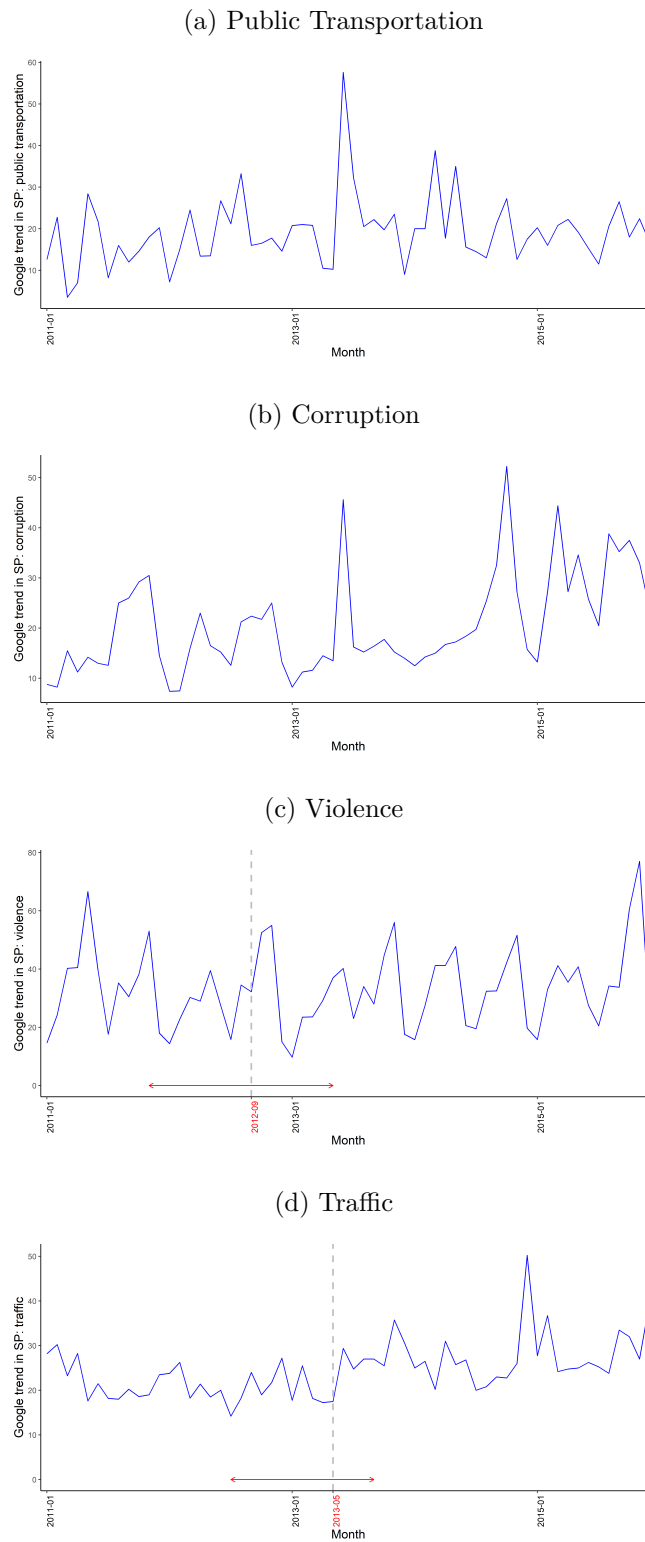
(a) Public Transportation

(b) Corruption

(c) Violence

(d) Traffic

Notes: These figures show the dynamic in search entries from Google Trends for each specific term unrelated to the drought. The gray vertical lines indicate the estimated structural break and the red dates on the x axis show the estimated date of the break. Red horizontal lines show the confidence interval for each break point calculated by the algorithm.

2.C   Spatial heterogeneities

Low pressure in the water distribution system might lead to stronger water shortages in regions with a higher altitude. Figure 2.G.4 shows that the MRSP has a heterogeneous ground elevation, therefore, we investigate if the altitude of a CT influences the impact of the pressure reduction policy on hospitalization by fecal-oral transmitted disease. It's important to notice that, as presented in the descriptive statistics from Table 2.3.1, the regions exposed to the pressure reduction policy are, on average, at a lower altitude. This is due to the fact that the Pressure Reduction Valves - used to define regions most likely exposed to pressure reduction policy - are usually installed on regions with a lower altitude.

Each column of Table 2.C.1 presents a sub-sample of CTs in different altitude groups, and the row "Mean altitude" shows the average altitude in each group of CTs. This heterogeneous analysis suggests that the altitude of a CT does not affect the impact of the policy on hospitalization.

Table 2.C.1 – Hospitalization by fecal-oral transmitted disease by average altitude of CTs

	Dependent variable			
	Altitude quartiles			
	(1)	(2)	(3)	(4)
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Panel A: low-income neighborhoods				
lowPress:dt	0.477	0.656**	0.188	0.164
	(0.323)	(0.278)	(0.242)	(0.248)
Mean dep. var	3.31	3.04	2.88	2.97
Mean altitude	747.43	771.04	788.02	819.72
Observations	118,860	130,920	147,540	185,940
Adjusted R^2	0.036	0.036	0.030	0.038
Panel B: high-income neighborhood				
lowPress:dt	−0.108	0.266	0.542***	0.196
	(0.260)	(0.177)	(0.175)	(0.260)
Mean dep. var	2.23	2.07	1.83	2.03
Mean altitude	747.59	770.61	787.64	813.11
Observations	187,020	173,040	158,160	118,200
Adjusted R^2	0.097	0.114	0.041	0.176

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Each column shows the results for a group of Census Tract in a given altitude group, where column 1 represents the lower group CTs and column 4 the higher. Each panel shows a different income group. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita.

The water provided by the water company must travel from the water towers to the households. It is possible that the longer the water travels inside the pipes, the higher the probability of infiltration from exterior contaminants. We also test this hypothesis by dividing the CTs according to its distance to the closest water tower. We calculated this distance by a proxy considering the euclidean distance between the centroid of each CT and the closest water tower. To locate the water

towers in the MRSP we georeferenced a map from the 2011 environmental impact report of the Production System São Lourenço that contained information about the water metropolitan system (SABESP, 2011). These results suggest that there is no direct relationship between the effect of the pressure reduction policy and the distance to a water tower.

Table 2.C.2 – Hospitalization by fecal-oral transmitted disease by average distance to the closest water tower

	Dependent variable			
	Distance to the closes water tower quartiles			
	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Panel A: low-income neighborhoods				
lowPress:dt	0.615**	0.407	0.534**	−0.008
	(0.295)	(0.283)	(0.251)	(0.225)
Mean dep. var	3.11	3.45	3.28	2.54
Mean altitude	793.99	786.58	784.21	783.01
Mean dist. water tower	0.68	1.35	1.95	3.7
Observations	104,280	124,200	150,540	204,240
Adjusted R^2	0.032	0.043	0.032	0.033
Panel B: high-income neighborhood				
lowPress:dt	−0.004	0.507***	0.146	0.057
	(0.168)	(0.170)	(0.223)	(0.246)
Mean dep. var	1.81	2.19	2.33	1.83
Mean altitude	780.74	775.01	773.09	772.62
Mean dist. water tower	0.67	1.34	1.92	3.07
Observations	200,640	180,720	154,380	100,680
Adjusted R^2	0.038	0.129	0.168	0.036

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Each column shows the results for a group of Census Tract dividing them according to the distance to the closest water tower. Column 1 show the results for CTs close to a water tower, while column 4 presents the results for CTs far from the towers. Each panel shows a different income group. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita.

2.D   Alternative specification

Table 2.D.1 shows different specifications for the main results. In columns 1 to 4, we do not include the level of each reservoir as a control, so we are able to use all CTs from the MRSP, but those within municipalities with few zip codes (in these initial models, we use all non-dark CTs from Figure 2.G.1). Column (1) is a simple difference-in-differences without any fixed effects or controls. In column (2) we add CT and period fixed effects. In columns (3) we interact the period fixed effect with a municipality fixed effect. In column (4) we further add a specific trend for each CT. In column (5) we reduced the sample to the preferred selection and run the model with CT fixed effect, a period by municipality fixed effect, but without controlling for the level of each reservoir. In column (6), which corresponds to the benchmark specification discussed in Section 2.4, we include control for the level of the reservoir. Finally, in column (7) we add a CT specific trend to the main specification.

Table 2.D.1 – Hospitalization by fecal-oral transmitted disease using different specifications

	Rate of hospitalization: fecal-oral						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: low-income neighborhoods							
lowPress:dt	0.282**	0.282**	0.217	0.389*	0.365**	0.361**	0.500**
	(0.121)	(0.124)	(0.140)	(0.228)	(0.151)	(0.149)	(0.234)
Mean dep. var	2.83				3.03		
Observations	805,740	805,740	805,740	805,740	583,260	583,260	583,260
Adjusted R^2	0.0004	0.030	0.033	0.029	0.033	0.033	0.029
Panel B: high-income neighborhood							
lowPress:dt	0.112	0.112	0.079	0.133	0.240**	0.192*	0.215
	(0.103)	(0.109)	(0.110)	(0.180)	(0.101)	(0.104)	(0.209)
Mean dep. var	2.02				2.05		
Observations	934,320	934,320	934,320	934,320	636,420	636,420	636,420
Adjusted R^2	0.00000	0.200	0.202	0.114	0.104	0.104	0.058
CT and period FEs		x	x	x	x	x	x
Mun. x period FE			x	x	x	x	x
CT spec. trend				x			x
Preferred sample					x	x	x
Lvl of reservs						x	x

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Dependent variable is the hospitalization by fecal-oral transmitted disease and each column uses a different specification

The results seem to be stable across the different models (especially in low-income neighborhoods), but we notice a slight increase in magnitude when using our preferred sample selection.

Table 2.D.2 repeats the bottled water price analysis, but calculating the price in each CT as an average of prices within 10 km of the CT. The results are weaker but similar to those observed in the previously presented model.

Table 2.D.2 – Observed effect in bottled water prices

	Dependent variable			
	Within 10 km price of bottled water			
	(1)	(2)	(3)	(4)
	Low-income		High-income	
lowPress:dt	−0.002	−0.004	−0.009***	−0.008***
	(0.008)	(0.010)	(0.003)	(0.003)
cant:dcant	0.0771**		−0.009	
	(0.036)		(0.017)	
dt:cant		0.095**		0.014
		(0.039)		(0.021)
lowPress:dt:cant		0.004		−0.005
		(0.008)		(0.006)
Mean dep. var.	1.87		1.88	
Observations	548,849	548,849	602,445	602,445
Adjusted R^2	0.702	0.703	0.777	0.777

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. The dependent variable is the average price of bottled water within 10 km of each CT. Columns 1 and 2 analyze the effects in low-income regions, while columns 3 and 4 are for high-income CTs. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita. Columns 1 and 3 use the specification from Equation 2.6.1, and columns 2 and 4 use specification from Equation 2.6.2.

2.E   Alternative event-study specification

This section presents the event-study for fecal-oral transmitted diseases using the monthly coefficients. Although the evidence is similar to what was observed in the previous exercise, in the monthly analysis the results are noisier.

Figure 2.E.1 – Monthly event-study for the hospitalization by fecal-oral diseases
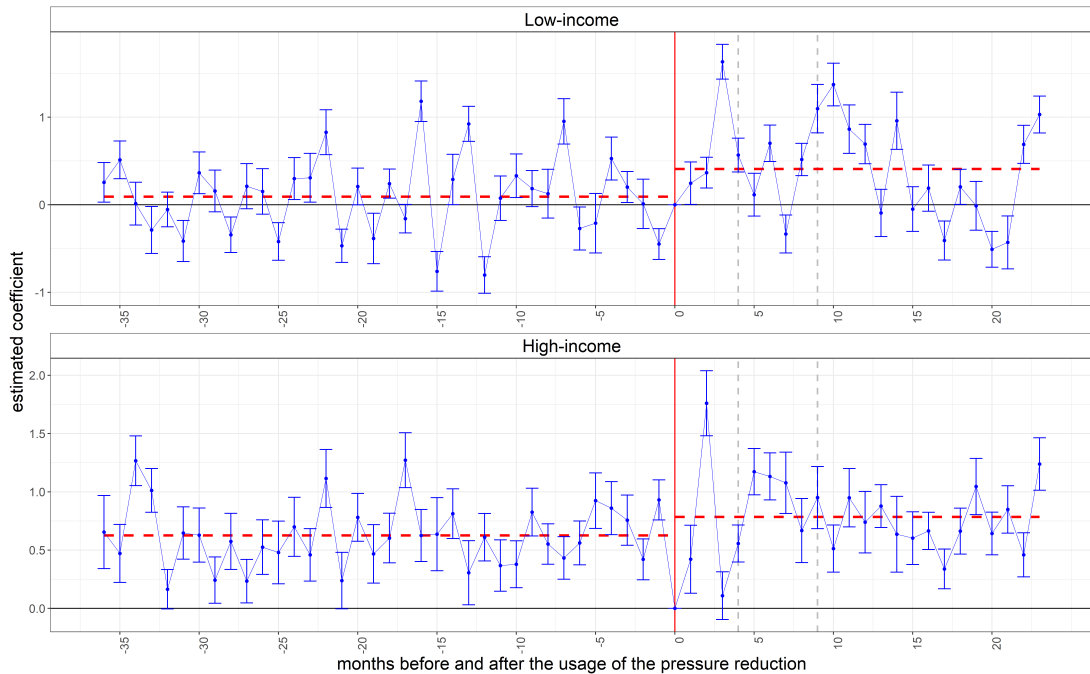
This figure shows the estimated coefficients for the event-study showing the impact of the pressure reduction policy on hospitalization in each month before and after the beginning of the policy.

2.F   Alternative spatial aggregation

This section repeats the main exercise of this study, but uses a larger spatial unit. We aggregate the Census Tracts into Weighting Area level (in Portuguese called *"Área de Ponderação"*). This is a regional aggregation, usually bigger than the Census Tract, used by IBGE to calibrate weights of each surveyed observation.[2.F.1]

We consider a Weighting Area to be within a specific reservoir if more than 50% of its area lies within that specific water system. The same rule was applied

---

[2.F.1]These Weighting Areas aggregate contiguous Census Tracts from the same district. This aggregation aims to group similar Census Tracts according to population and known infrastructure characteristics (including share of private residential households, share of households connected to water and sewage systems, average number of residents per dwellings, among others). At the 2010 IBGE Census, Brazil was divided into 10,184 Weighting Area and 316,574 Census Tracts (IBGE, 2010).

to define a region exposed to the pressure reduction policy. Figure 2.F.1 shows the map of the MRSP with blue regions considered to be affected by the policy.

Figure 2.F.1 – Map of pressure reduction area considering a larger spatial aggregation
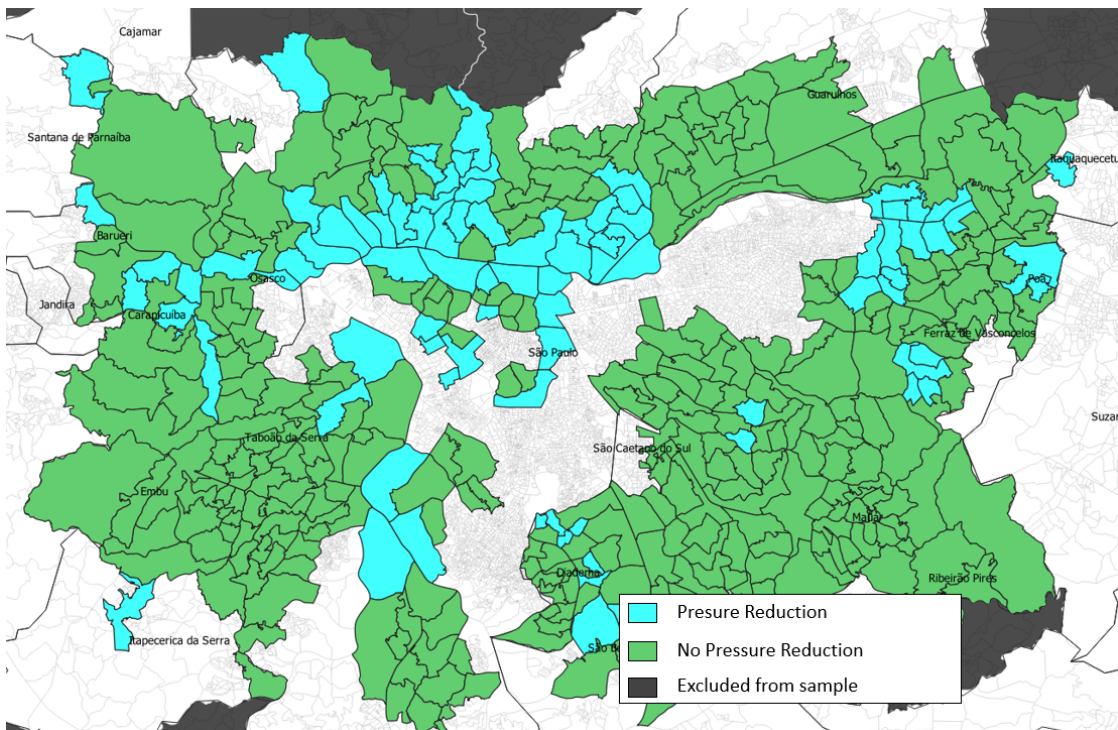
Notes: Figure showing the main sample aggregated to the level of a Weighting Area, as defined by IBGE.

Table 2.F.1 presents the results from this exercise, that corroborate with our main conclusions, but loose statistical significance. This might be because now we have fewer observations in our sample.

Table 2.F.1 – Aggregating to the Weighting Area level

	Hospitalization by fecal-oral disease			
	Rate			Level
	(1)	(2)	(3)	(4)
Panel A: low-income neighborhoods				
lowPress:dt	0.424**	0.558**	0.575*	0.193
	(0.213)	(0.252)	(0.305)	(0.123)
Mean dep. var	2.85	3.04	3.04	1.15
Observations	19,320	12,960	12,900	12,900
Adjusted R^2	0.007	0.008	0.234	0.303
Panel B: high-income neighborhood				
lowPress:dt	0.138	0.161	0.008	0.0001
	(0.122)	(0.150)	(0.153)	(0.054)
Mean dep. var	1.98	2	2	0.69
Observations	17,160	11,400	11,280	11,280
Adjusted R^2	0.0004	0.002	0.249	0.184
Preferred sample		x	x	x
CT FE			x	x
Period x Mun FE			x	x
Level control			x	x

Notes: *p<0.1; **p<0.05; ***p<0.01. Two-way clusters at the Census Tract and date levels. Robust standard errors are in parentheses. Regressions are weighted by the number of inhabitants in each CT. Columns 1 through 3 use the rate of hospitalization by a hundred thousand inhabitants as the dependent variable, while column 4 uses hospitalization in level. Low-income Census Tracts are defined as CTs where more than half of the households live with less than one minimum wage per capita.

2.G   Maps library

Figure 2.G.1 – Full MRSP
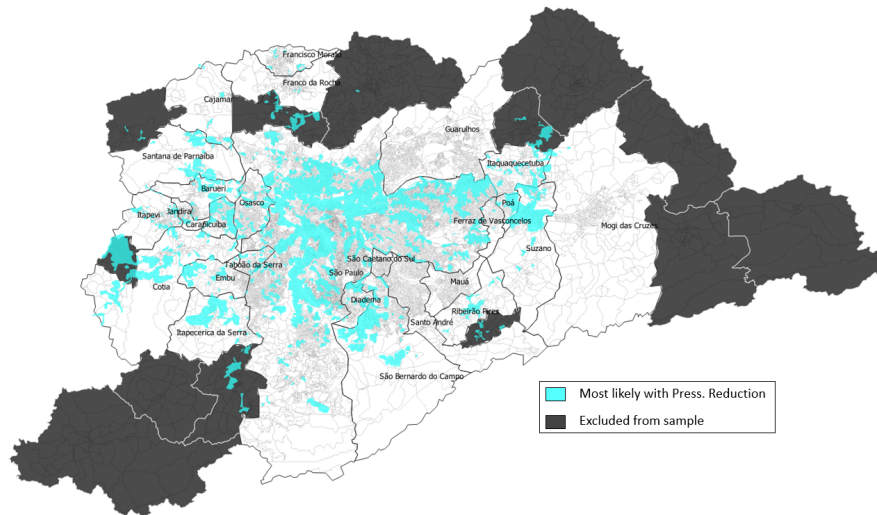
Notes: Blue regions are the most likely exposed to pressure reduction; black regions are municipalities from the MRSP excluded from the analysis.

Figure 2.G.2 – Main sample
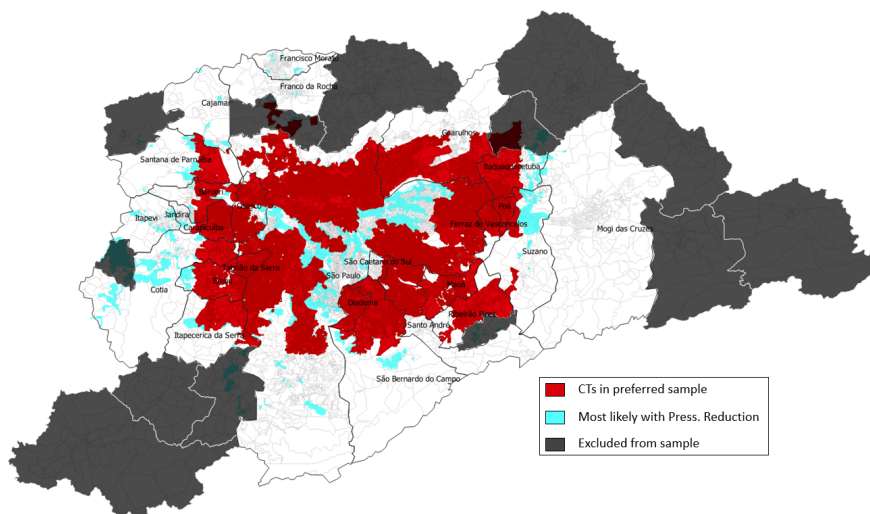
Notes: This figure shows all the MRSP highlighting the main sample from the analysis in red.

Figure 2.G.3 – Regions supplied by each water reservoir used in this study
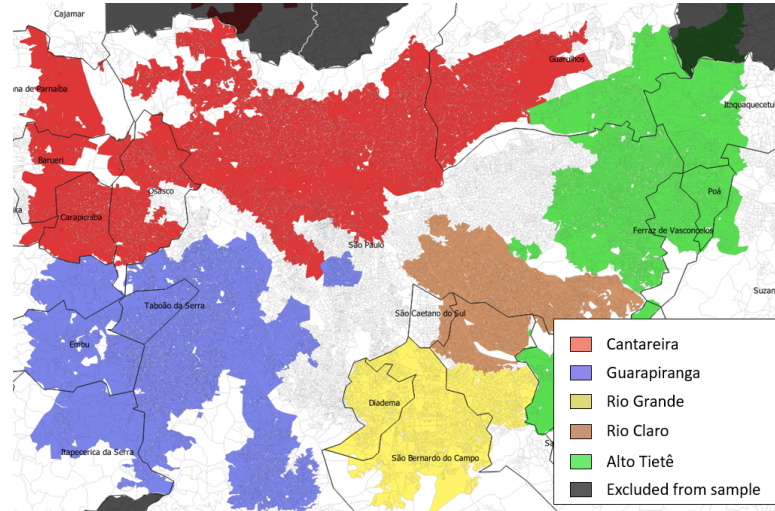
Notes: Regions considered to be supplied by each of the water reservoirs throughout the drought period.

Figure 2.G.4 – Average altitude of Census Tracts
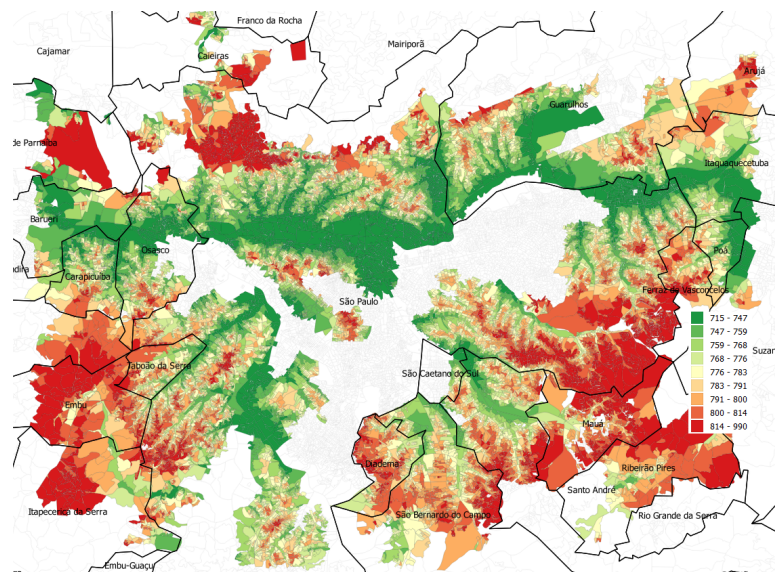
Notes: This figures, built by the authors with data from Topodata project from INPE, shows the average altitude of each Census tract from the main sample.

3 Health Consequences of Thermoelectric Power Plants

3.1 Introduction

Thermoelectric power plants are notorious stationary sources of air pollution, but many countries still rely on them as an important source of energy. The role of this type of power on the emission of Greenhouse Gas (GHG) is largely discussed by scientists due to its impacts on the ongoing process of climate change. Still, increasingly more attention is paid to the short and long-term impacts of the energy sector on local air pollution and its health consequences. Because of its high emissions of carbon dioxide, an important GHG, power plants fueled by coal are usually at the center of the discussion, but, despite not being major threats to climate change, other fuels might also generate local air pollution with deleterious consequences to the human health.

In this study, we analyze the health consequences of different types of thermoelectric power plants in Brazil, a region with a low share of thermoelectric power on its energy mix. The country relies mostly on renewable energy sources, with hydroelectric power accounting for 65% of the generation and almost 7% of wind and solar power. Despite this fact, thermoelectric power has increased in the past few years, up from 15% in 2009 to around 20% in 2017. This increase is mainly driven by natural gas-fueled power plants (EPE, 2018a; EPE, 2010).[3.1.1] In the past few decades, Brazil has also drastically reduced the mortality rate of children under five years of age, which changed from 53 deaths by a thousand live births in 1990 to 13 in 2018 (UNICEF, 2019). Since air pollution can potentially have adverse effects on health outcomes, it is important to understand if the use of thermoelectric power in Brazil can affect the path of further improving health in early childhood and how the different types of fuel influence the health conditions of the surrounding population.

The connection between air pollution and health is widely discussed in the

---

[3.1.1]See Figure 3.2.1 for more details on the Brazilian energy mix.

literature, and there is evidence pointing to a positive relation.[3.1.2] While most of the literature considers a developed country, some studies focus on developing countries, where there is usually a challenge to assess the data related either to mortality or the air pollution levels.[3.1.3] Arceo, Hanna e Oliva (2016) suggest that the observed effect of pollution on health in developing countries might differ from what the literature usually finds in developed countries. They argue that a non-linear dose-response relationship between pollution and infant mortality could overly harm developing countries that usually observe higher levels of pollution. Besides that, they point that behavioral effects could also differ between these groups of countries. Although infant mortality is the most dramatic consequence of air pollution, many analyses also investigate its effects on pregnancy quality as reflected by health outcomes after birth.[3.1.4]

The Fetal Origin Hypothesis, discussed by Almond e Currie (2011) and Currie (2009), suggests that health shocks before birth can influence people's health throughout the whole life.[3.1.5] Isen, Rossin-Slater e Walker (2017) analyze such claims using the enactment of the Clean Air Act in 1970 in the United States as a source of exogenous variation in air pollution. They find evidence that newborns more exposed to pollution in the year of birth presented worse labor market outcomes some 30 years later. There is also evidence that the consequences of pollution shocks during pregnancy have heterogeneous effects according to gender and the type of pollutant. Sanders e Stoecker (2015) and Jedrychowski et al. (2009) find that male children are more vulnerable to air quality shocks before birth, but the opposite result has also been found (TANAKA, 2015). A constant concern from the literature analyzing birth outcomes and infant deaths is that the observed adverse effects of shocks during pregnancy are understated due to fetal selection.

[3.1.2]Deryugina et al. (2019), Knittel, Miller e Sanders (2016), He, Fan e Zhou (2016), Tanaka (2015), Sanders e Stoecker (2015), Greenstone e Hanna (2014), Luechinger (2014), Jayachandran (2009), Currie, Neidell e Schmieder (2009), Currie e Neidell (2005), Chay e Greenstone (2003)

[3.1.3]Tanaka (2015), Greenstone e Hanna (2014), Jayachandran (2009), and He, Fan e Zhou (2016).

[3.1.4]Rangel e Vogl (2019), Gehrsitz (2017), Yang e Chou (2018), Knittel, Miller e Sanders (2016), Currie e Walker (2011).

[3.1.5]There are many types of shocks that could affect pregnancy, birth outcomes, and the future well being of an individual, such as violence (KOPPENSTEINER; MANACORDA, 2016), stress (PERSSON; ROSSIN-SLATER, 2018), water scarcity (ROCHA; SOARES, 2015), among others.

This selection happens because the weaker fetuses die before ever being born, so the sample of newborns observed by the researchers is a selection of the strongest individuals. Because only the stronger fetuses survive, the fetal selection might induce the researchers to underestimate the real negative consequences of the shocks during pregnancy on birth outcomes.[3.1.6]

Hospital admission is another variable usually analyzed by researchers interested in investigating the health consequences of air pollution.[3.1.7] This variable allows the researcher to evaluate the consequences of pollution on distinct stages of life. However, results usually suggest that the most affected groups are the children and the elderly population. Air pollution also has consequences that go beyond a worsening in health outcomes, and the literature has analyzed its impacts on labor productivity (HANNA; OLIVA, 2015; Graff Zivin; NEIDELL, 2012), education (CHEN; GUO; HUANG, 2018), crime (BURKHARDT et al., 2019), among others.

A challenge from the literature that investigates the impact of pollution on health is to isolate the effect of air pollution from other determinants of health. Many studies try to isolate this effect by either finding an adequate instrument or rely on a quasi-experiment. Some strategies adopted in recent works are to use: changes in regulation and in the economic activity;[3.1.8] weather instruments and variations in the wind direction;[3.1.9] exogenous changes in large emission sources, such as ports, airport, agricultural fires, and stationary pollutant facilities.[3.1.10]

Changes in the energy sector are also used as exogenous variations in some works from this literature. Luechinger (2014) analyzes a German policy aiming to install scrubbers to reduce Sulfur Dioxide(SO_2) emissions from power plants. The author uses this new regulation, along with predominant wind direction in

---

[3.1.6]Rocha e Soares (2015), Currie (2009).

[3.1.7]Guidetti, Pereda e Severnini (2020), Schlenker e Walker (2016), Chagas, Azzoni e Almeida (2016), Moretti e Neidell (2011), Neidell (2009).

[3.1.8]Yang e Chou (2018), Gehrsitz (2017), He, Fan e Zhou (2016), Sanders e Stoecker (2015), Tanaka (2015), Greenstone e Hanna (2014), Luechinger (2014), Chay e Greenstone (2003).

[3.1.9]Guidetti, Pereda e Severnini (2020), Deryugina et al. (2019), Rangel e Vogl (2019), Arceo, Hanna e Oliva (2016).

[3.1.10]Rangel e Vogl (2019), Yang e Chou (2018), Severnini (2017), Schlenker e Walker (2016), Knittel, Miller e Sanders (2016), Chagas, Azzoni e Almeida (2016), Hanna e Oliva (2015), Moretti e Neidell (2011), Currie e Walker (2011), Jayachandran (2009), López et al. (2005).

the power plants, to isolate the effect of SO_2 concentration on infant mortality. Results point to a significant reduction in mortality and suggest that the policy avoided around a thousand infant deaths a year. Severnini (2017) shows that the shutdown of nuclear power plants in the Tennessee Valley led to an increase in the use of coal-fueled power plants, which increased pollution in counties where the facilities were located. The author finds that this increase in the concentration of Total Suspended Particle (TSP) had a negative effect on the average birth weight of newborns. Yang e Chou (2018) evaluate that the shutdown of a coal-fired power plant led to a reduction in the probability of having low-birth weight and preterm-born babies in the regions downwind of the facility. Ordonez (2020) uses river flows as an instrument for thermoelectric power generation and investigates the effect of fossil fuel facilities on pollution concentration and hospitalization. His results suggest that the use of thermoelectric power plants can lead to issues in respiratory and mental health. His analysis also suggests that diesel and coal facilities have the most significant effect on pollution, with no observed negative consequence of natural gas power plants.

Most of the literature analyzing the effects of power plants on pollution and health investigate the consequences of coal-fueled facilities, because they account for a large share of the world energy production. Although considered cleaner, other types of power plants are also sources of air pollution and can influence the health of people living nearby. Ordonez (2020) showed that different fuels had distinct impacts on pollution and on hospitalization. We further contribute to the literature by investigating the effects of the different types of power plants from a developing country on the infant mortality rate. We use wind direction and the schedule of energy generation from different thermoelectric facilities in Brazil to investigate the health consequences of exposure to a nearby operational power plant. We build a variable that captures exposure to an operational power plant during pregnancy and investigate its effects on infant mortality.

The main results from this analysis suggest that exposure to coal power plants leads to a small increase in the infant mortality rate. Although the result is consistent across different specifications, the magnitude of the effect is small. Every additional Gigawatt hour (Gwh) of energy generated that influences a municipality

leads to an increase in 0.003 deaths in the first year of life by a thousand live births. The results are robust in suggesting that the infant mortality effect is mainly driven by female children. We also find evidence that the population living closer to the power plants are more affected by it and that fetal selection is likely to interfere with the observed results biasing them toward zero. To have a broader sense of the health consequences of using thermoelectric power, we also investigate the effect of the contemporaneous exposure to power plants on hospitalization by air-related diseases in different age groups, but there is no evidence that increasing the exposure in a given month leads to an increase in hospitalization in the same period.

This paper is organized as follows. The next section discusses the energy market in Brazil. Section 3.3 presents the data used in this analysis. In Section 3.4, we discuss the methodology. Next, in Section 3.5, we present the main results, and Section 3.6 concludes this work.

3.2   Brazilian energy market

Brazil is a country that heavily relies on hydroelectric power plants to generate its energy, as noted in Figure 3.2.1. In the country's regulatory framework, a centralized agency called National System Operator (ONS) coordinates energy production from every agent supplying for the National Interconnected System (SIN), which encompasses most of the energy produced in the country. The system is divided into four subsystems (North, Northeast, Southeast/Central-West, and South) that are connected to the same grid. ONS is responsible for optimally dispatching energy from different sources and regions, considering the trade-offs between current and future water availability in reservoirs and the energy production cost.

Figure 3.2.1 shows that in recent years the use of wind power, and especially the use of thermoelectric power plants, has been increasing. Changes in the hydrological cycle contributed to this shift in energy sources (MENDES; STHEL, 2017), and this movement away from hydropower is likely to keep happening in the

next few years (EPE, 2018b, p. 67). The left panel of Figure 3.2.2 points out that the increase in thermoelectric power taking place in the country is mainly focused on natural gas power plants, while the right panel shows that many large power plants were built only in the last decades.

Figure 3.2.1 – Brazilian energy mix
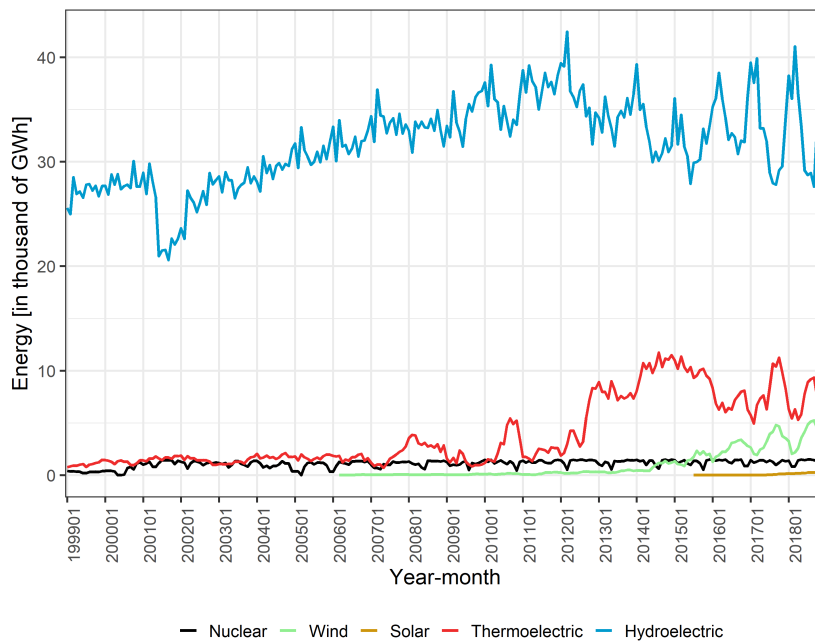
Notes: Figure build by the author with information taken from ONS. This figure shows the monthly generation by source of energy.

Figure 3.2.2 – Increase in the use of thermoelectric power plants

Besides being a larger source of carbon dioxide (CO_2), that directly contributes to climate change, thermoelectric power plants also generate a series of other pollutants that can be harmful to human life. The quantity and type of pollutant depend on the technology and type of fuel used. Particle matters (PM) are small particles and liquid droplets that, if inhaled, can harm human health. They are emitted by natural and human-induced processes such as dry soil, pollen, sea salt spray, agricultural fires, some industrial facilities, power plants, among others. Particles with a diameter under 2.5 micrometers are referred to as $PM_{2.5}$ and are among the most harmful types of air pollution (MULLER; MENDELSOHN, 2007).

Carbon monoxide (CO) is a colorless and odorless toxic gas generated by the incomplete combustion of fuel that can harm the absorption of oxygen (O_2) by the respiratory system, with known negative consequences for human health (CURRIE; NEIDELL, 2005; CURRIE; NEIDELL; SCHMIEDER, 2009). In our society, CO is mainly produced by the transport sector, but thermoelectric power plants can also generate it (HINRICHS; KLEINBACH, 2012). Sulfur oxides (SO_x) are primarily generated by the use of fossil fuel power plants, and there is evidence that it has adverse effects on infant health (LUECHINGER, 2014). Oxidation of sulfur dioxide generates SO_3 that, when reacting with water, becomes sulfuric acid that can lead to acid rain. This phenomenon has severe consequences for the

environment, the economy, and the health of affected communities.

Volatile Organic Compounds ($VOCs$) are compounds of carbon and hydrogen atoms. They are produced by natural and human activities, such as fuel combustion, evaporation of industrial solvents, and biological decay of vegetation. Most of the $VOCs$ are generated by natural sources, but those emitted by human activities are usually the more reactive ones. Nitrogen-oxygen compounds (NO_x) are mainly generated by the transport sector, but thermoelectric facilities can also lead to their formation in the atmosphere. This sort of pollutant is not generated by the burned fuel itself, but by the air used in the combustion process (atmospheric air contains around 78% of nitrogen). NO_x reacts with $VOCs$ in the presence of light to generate ozone, which is a major source of air-related health problems, especially for people used to practice outdoor activities (NEIDELL, 2009).

The impact of a power plant on air quality depends on many factors like the type of fuel, the type, and age of the generation and control equipment, the location, local weather, among others (MASSETTI et al., 2017). Coal, natural gas, and oil are all fossil fuels used to power a thermoelectric facility. The level of emissions from coal-fired power plants depends on the quality of the burned coal. Significant pollutants in this source of energy are carbon dioxide (CO_2), nitrogen oxides (NO_x), sulfur oxides (SO_x), and particle matters (PM). In Brazil, the emission rates of SO_2 and PM are higher in power plants using national coal than in those that use the imported fuel (EPE, 2007b).

Natural gas power plants are usually a cleaner source of energy, emitting less CO_2, but NO_x and CO can also be observed in this type of facility. A relative advantage of natural gas compared to other fossil fuels is its low emission of PM and SO_x (TOLMASQUIM, 2016). For the oil-fueled power plants, the main emitted pollutants are CO_2, NO_x, and SO_2, but, depending on the fuel's characteristics, smaller amounts of CO can also be observed (EPE, 2007b; AVELINO; HEWINGS; GUILHOTO, 2014).

Biomass power plants can be fired by different types of fuels, some of which are residuals and byproducts of industrial or agricultural processes. Still, some commodities are grown specifically to be used as fuels for the power plants. Biomass-

fueled power plants are considered renewable sources of energy, and, in Brazil, they are usually fueled by sugar cane. Because they are associated with the sugar cane growth cycle, this type of energy has a seasonal pattern, and it is usually active between April and October (TOLMASQUIM, 2016). Particulate matter (PM) is the most significant emission from biomass power plants, but NO_x and CO can also be observed depending on the specific fuel used (EPE, 2007a).

In Brazil, the central government, through the National Board for the Environment (CONAMA - *"Conselho Nacional do Meio Ambiente"*), sets minimal air quality standards for the country, but local authorities can define stricter rules. The first national regulation is the CONAMA resolution number 3, from 1990 (CONAMA, 1990). This resolution defined a threshold for maximum concentration of Total Suspended Particles, smoke, SO_2, CO, O_3, and NO_2. Although an important sign to indicate that the country was perusing better air quality standards, this regulation set lax thresholds for pollution concentration when compared to the current best practices, and there were no specific rules for $PM_{2.5}$. In 2018, after almost 30 years of the original regulation, a new resolution from CONAMA (Resolution 491/2018) established new standards in line with the World Health Organization's guidelines (WHO, 2006), which includes rules for the concentration of $PM_{2.5}$. Although the objectives of this new regulation respect WHO standards, Siciliano et al. (2020) points that these standards are defined as a goal to be reached in the future but without a clear road map of implementation, as each State will structure its own plans.

3.3   Data

Assessing pregnant mothers' and newborns' exposure to pollution is a challenge from this literature because the researcher rarely observes the amount of pollution inhaled by each person.[3.3.1] Usually, economic studies consider the readings of pollution from monitoring stations near the population of interest (zip

---

[3.3.1]There are works from the medical literature that give portable air quality sensors to the research subjects, but even this type of work is usually done to a limited number of individuals and covers only a limited time period (JEDRYCHOWSKI et al., 2009).

code areas or municipalities). Unfortunately, Brazil does not have a large network of air quality monitoring stations spread throughout the country, therefore, we use a reduced form strategy to investigate if an operational power plant affects the downwind population's health outcomes.

3.3.1 Health outcomes

The health variables used throughout this study are gathered by the Brazilian Ministry of Health. The three data sets used are the System of Mortality Records (SIM/Datasus), the System of Live Birth Records (SINASC/Datasus), and the System of Hospitalization Records (SIH/Datasus).

The SIM provides data on every death officially registered in the country containing information such as the age (and date of birth), the municipality of residence, the cause of death, the gender of the dead, etc. As the main dependent variable, we select every infant's death on the first day, first week, first month, first three months, and first year of life. This mortality data is aggregated at the month of birth and municipality level and the spanning period of the analysis goes from January 2010 to December 2016. Data from 2018 are still preliminary at Datasus, so we will not use it. Because the main dependent variable is death in the first year of life, the last month in the sample has to be December 2016 so we can use complete information up to December 2017. Figure 3.A.1 shows that the infant mortality rate has been fairly stable in the past few years, around 12 deaths per thousand live births.

We will also use data from hospitalization to investigate the broader effects of power plants on health. The SIH contains information about every hospitalization paid for by the Brazilian public health system. It informs the patient's zip code and municipality of residence, the date of the hospital admission, and the cause of it. It also shows the age and gender of every patient, along with other information. We will aggregate this data to the municipality by month level to analyze if a month with high exposure to power plant influences the hospitalization rate in different age groups.

The mortality and hospitalization data contain information about the main diseases that led to the health issue. We selected disease groups that are likely to be influenced by air pollution to investigate their prevalence on the population exposed to power plants. We will divide the causes of infant mortality into three main groups that account for a large share of the infant deaths and have some evidence that might be related to air pollution (VRIJHEID et al., 2011; TEIXEIRA et al., 2019): respiratory illnesses; conditions that originated in the perinatal period; and congenital malformation.[3.3.2] We also analyze subgroups of respiratory diseases: Asthma (ICD-10 codes J45-46), Acute Respiratory diseases (ICD-10 codes J00-06, J20-22), Chronic Obstructive Pulmonary Disease - COPD (ICD-10 codes J40-44), and Pneumonia (ICD-10 codes J12-18).[3.3.3]

Finally, we use birth data to analyze health at birth and to calculate the number of live births in each municipality and month. SINASC from Datasus contains data of every live birth in Brazil, informing the length of the pregnancy (grouped in coarse week periods), birth weight, APGAR1, APGAR5, date of birth, municipality of residence, previous pregnancies of the mother, type of birth (cesarean or natural birth), prenatal care (how many times has the mother seen a doctor during pregnancy), child's gender, among other variables. We aggregate this data at the month of birth by municipality level. The infant mortality rate used throughout this study is calculated using data from SIM and SINAC. It considers deaths in the first year of life of children born in a given month (from SIM), divided by the a thousand live births in that month (from SINAC). The descriptive statistics of the main health variables are presented in Table 3.3.1.

[3.3.2]Chapters X (Diseases of the respiratory system), XVI (Certain conditions originating in the perinatal period), and XVII (Congenital malformations, deformations and chromosomal abnormalities) of the International Classification of Diseases (ICD). Selected diseases for these groups are those whose ICD codes starts with the letters "J", "P", and "Q" respectively.

[3.3.3]Disease groups defined similar to Tao et al. (2014) and Song et al. (2018).

Table 3.3.1 – Health characteristics of the main sample

Variable	Mean	SD	Min	Max
IMR (deaths under 1 year of age by a thousand live births)	12.12	13.23	0.00	1000.00
Male IMR	6.66	9.74	0.00	1000.00
Female IMR	5.41	8.76	0.00	1000.00
IMR from resp. diseases	0.63	2.74	0.00	1000.00
IMR Perinatal	6.96	10.17	0.00	1000.00
IMR malform.	2.73	6.09	0.00	1000.00
Rate of asthma: under 1	4.35	11.42	0.00	542.89
Rate of acute: under 1	21.81	35.24	0.00	855.80
Rate of COPD: under 1	1.74	6.97	0.00	630.25
Rate of pneumonia: under 1	36.20	45.89	0.00	1551.35
Rate of all resp: under 1	69.60	73.74	0.00	1861.62
Average live births	75.69	116.77	0.00	682.00
Birth weight	3181.36	81.56	466.00	5400.00
Share low birth weight (< 2500)	0.09	0.03	0.00	1.00
Share very low birth weight (< 1500)	0.01	0.01	0.00	1.00

Notes: This table, built by the authors with data from Datasus, presents the descriptive statistics related to 500 munici-palities from the main sample.

3.3.2 Weather variables

Weather conditions are obtained from the *ERA5* reanalysis data, gathered by the ECMWF (European Center for Medium-Range Weather Forecasts). This dataset contains information on hourly atmospheric conditions with a resolution of around 30 km x 30 km.[3.3.4] To aggregate the grid data to the municipality level, we use a weighted average from each grid cell intersecting a municipality. The weights are calculated according to the share of the municipalities' land covered by each grid (similar to Auffhammer et al. (2013)). The weather control variables are given in monthly averages, but we also test an alternative definition for the temperature variable by dividing the daily averages into five bins and counting the number of times within a month that the daily average temperature falls into each bin (similar to Deschênes, Greenstone e Guryan (2009)).[3.3.5] The results are indifferent from the definition of weather control, as will be shown in Appendix Table 3.B.1. Throughout this study, we will use the variables as monthly averages

---

[3.3.4]https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

[3.3.5]The temperature bins are: under 10°C, between 10°C and 20°C, between 20°C and 30°C, between 30°C and 40°C, and above 40°C.

instead of bins. Table 3.3.2 presents the descriptive statistics of the municipalities' characteristics, as well as weather variables.

Table 3.3.2 – Municipalities' characteristics and weather variables

Variable	Mean	SD	Min	Max
Population [thousand]	56.59	83.54	1.56	458.67
Population: until 5 years old [K]	5.01	7.21	0.08	36.54
Population: 6 to 64 years old [K]	47.76	70.93	1.26	387.87
Population: above 65 years old [K]	3.68	5.75	0.15	58.91
Income per capita [2010 K BRL]	0.45	0.21	0.10	1.57
Nearby power plants	4.25	2.69	1.00	9.00
Average temperature	23.16	3.82	9.33	31.01
Average wind speed	2.46	1.10	0.52	7.61
Average daily precipitation	3.59	3.05	0.00	22.18

Notes: Descriptive statistics from municipalities' characteristics and weather conditions considering only the main sample. The information built by the authors with data from the 2010 Brazilian Census and from ERA5 weather data.

3.3.3 Exposure to energy generation

The main explanatory variable used throughout this analysis considers how much energy is produced in the vicinity of a municipality, as well as the population exposed to it. To build this variable, we need to know how much energy a power plant produces and verify to which municipalities the wind is transporting the generated pollution. Many highly complex models have been used to evaluate pollution dispersion and health outcomes[3.3.6], but simpler structures using wind direction, wind speed, and distance to the pollution source have also been successfully used.[3.3.7] A common strategy in this literature is to collapse the municipality to its centroid (or population-weighted centroid), evaluate if this point is within a range of the wind direction coming from the pollution source, and then verify the health outcomes in the municipality.

We use a similar strategy but also consider the share of the population in each municipality exposed to the pollution from the power plant. Figure 3.3.1 helps to illustrate how we build the exposure variable. Initially, we calculate the wind

---

[3.3.6] Avelino, Hewings e Guilhoto (2014), López et al. (2005).

[3.3.7] Rangel e Vogl (2019), Yang e Chou (2018), Schlenker e Walker (2016), Currie e Walker (2011).

direction in each power plant with data from ERA5 (the same data presented in Section 3.3.2). The direction in a power plant is based on the values observed in the four closest grid points. We calculate the average wind direction weighted by the inverse of the distance between the power plant and each of the four grid points. We aggregate the hourly wind direction from the four points to each power plant, and then calculate the daily averages. Panel A of Figure 3.3.1 shows a histogram of the daily wind direction in each octant of the wind rose at the Jorge Lacerda power plant in each month of the year. The further away from the center, the higher the frequency of the wind direction falling in that octant. The colors indicate the observed daily wind speeds.

Then, we use Census Tracts (CTs) from the 2010 Brazilian Census to evaluate the share of the population from a municipality exposed to each power plant in a given day. First, we calculate if the centroid of each CT from a municipality is downwind from a power plant in a given day. We consider the centroid of a CT to be downwind of a power plant if it is within 45° from the wind blowing from that power plant.[3.3.8] This process is illustrated in Panel B of Figure 3.3.1, which shows the example of the municipalities of *Capivari de Baixo* and *Tubarão*, that are close to the *Jorge Lacerda* power plants.[3.3.9] Suppose the wind is blowing in the East-West direction. CTs in green are considered to be downwind of the power plants, so we sum their population and divide it by the number of inhabitants of the entire municipality.

Finally, we multiply the share of the population exposed by the quantity of energy generated by the power plant. We use the information of the operation history of the electric system gathered by ONS (National System Operator), which contains data on the daily energy generation of each power plant. Panel C of Figure 3.3.1 shows the histogram of daily energy generation of the Jorge Lacerda power plant.

[3.3.8]If the absolute value of the difference between the wind direction and the power plant-census tract centroid direction is smaller than 22.5°, than the Census Tract is considered to be downwind of the power plant.

[3.3.9]There are three Jorge Lacerda power plants, but we are only showing the example for one of them.

Figure 3.3.1 – Exposure variable: example of Jorge Lacerda power plant

(a) Wind direction histogram
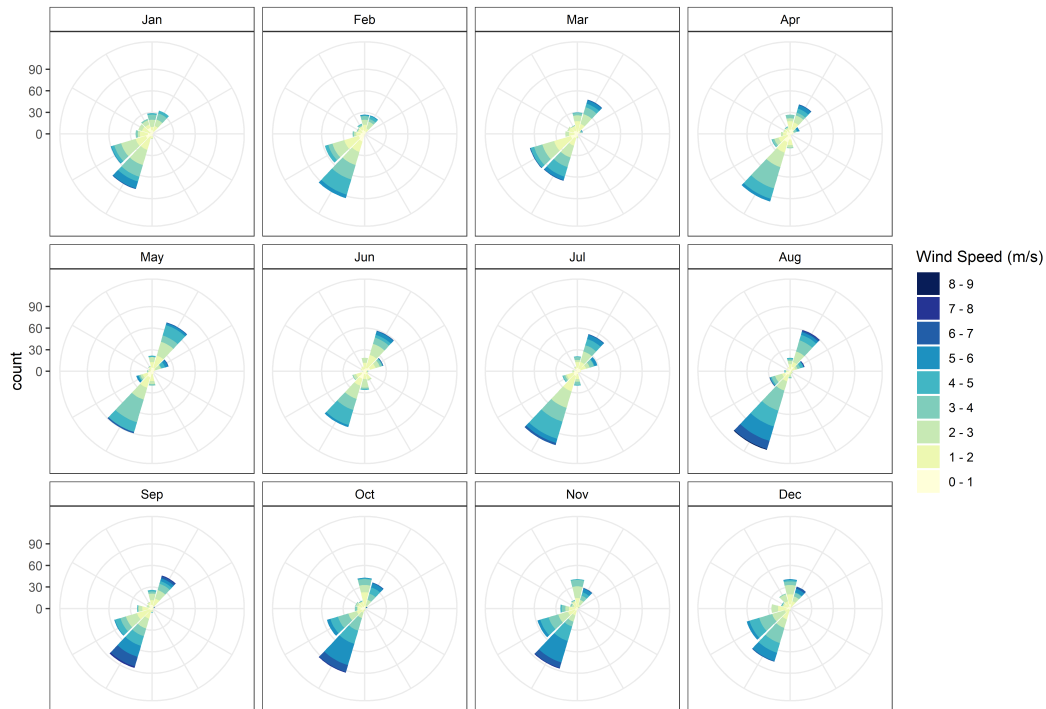
(b) Exposed Census Tracts

(c) Energy generation histogram
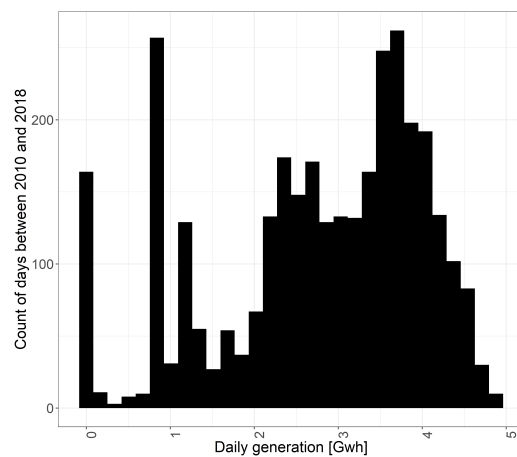
Notes: These figures show how the exposure variable is built for the particular case of the Jorge Lacerda A2 power plant, near the municipalities of *Capivari de Baixo* and *Tubarão*. Panel A presents a histogram of the frequency of the observed wind direction and speed. Panel B shows which Census Tracts are considered to be exposed to the pollution when the wind is blowing in the East-West direction. Panel C shows a histogram of the daily energy generation for the power plant.

82

This process calculates the exposure of one municipality to one power plant, but we use the same procedures to calculate the daily exposure of every municipality to every nearby power plant. Formally, this daily exposure variable is calculated as follows:

$$E_{id}^f = \sum_{\forall j \in i} \sum_{\forall s \in i} G_{jd}^f * W_{sjd}^i * p_{s,i} \tag{3.3.1}$$

where G_{jd}^f is the total energy generated on day d at thermoelectric power plant j fueled by f. W_{sjd}^i is an indicative variable equals 1 if the centroid of the Census Tract s from municipality i is downwind of the power plant j on day d. $p_{s,i}$ is the share of the population from municipality i that lives at Census Tract s, as recorded by 2010 Census. We sum over all Census Tracts (s) from municipality i and then over all power plants (j) within 50 km of the center of that municipality. The same municipality and CT can be exposed to different power plants, and the final exposure variable sums the exposure to each individual power plant.[3.3.10] This variable is the daily exposure of a municipality to operational power plants and it measures the amount of energy generated as well as the share of the population that was exposed to it. Because we work at the monthly level, we sum the daily exposure to find the monthly variable and then sum it with the eight months prior to the birth month m, as shown in Equation 3.3.2.

$$E_{im}^f = \sum_{\tau=m-8}^{m} \sum_{\forall d \in \tau} E_{id}^f \tag{3.3.2}$$

where τ assumes the value of each of the pregnancy months of a child born at month m, and d is each day from month τ.

Besides the data sources presented above, to build the exposure variable we also need data on thermoelectric power plants' locations and the type of fuel each of them use. The data gathered by ONS contains information about the daily energy generation for more than 200 thermoelectric power plants operating in the

---

[3.3.10]In Figure 3.3.1, we only presented the exposure to one of the Jorge Lacerda power plants, but we do the same thing for the other two, and then sum the results.

National Interconnected System, but around 95% of all the thermoelectric energy from this system comes from 75 facilities, so we focused this study on this group of the largest power plants. To find the fuel and location of these facilities, we used three main sources: (1) maps from the OpenStreetMaps, that contains user-inserted information about sites all around the world;[3.3.11] (2) the SIGEL dataset, provided by the Brazilian Electricity Regulatory Agency (ANEEL), that contains georeferenced information of power plants and also the type of fuel used;[3.3.12] and (3) articles from newspapers and the website of the companies. We couldn't find information about 5 facilities from the largest 75, so the final sample contains 70 thermoelectric power plants spread throughout the country, as presented in Figure 3.3.2.

[3.3.11]https://www.openstreetmap.org/

[3.3.12]Not all power plants from ONS appear in SIGEL's shapefile, and some facilities were placed in the center of a municipality, instead of its real location outside the urban area. That's why we also used other sources of information.

Figure 3.3.2 – Map of power plants from the sample
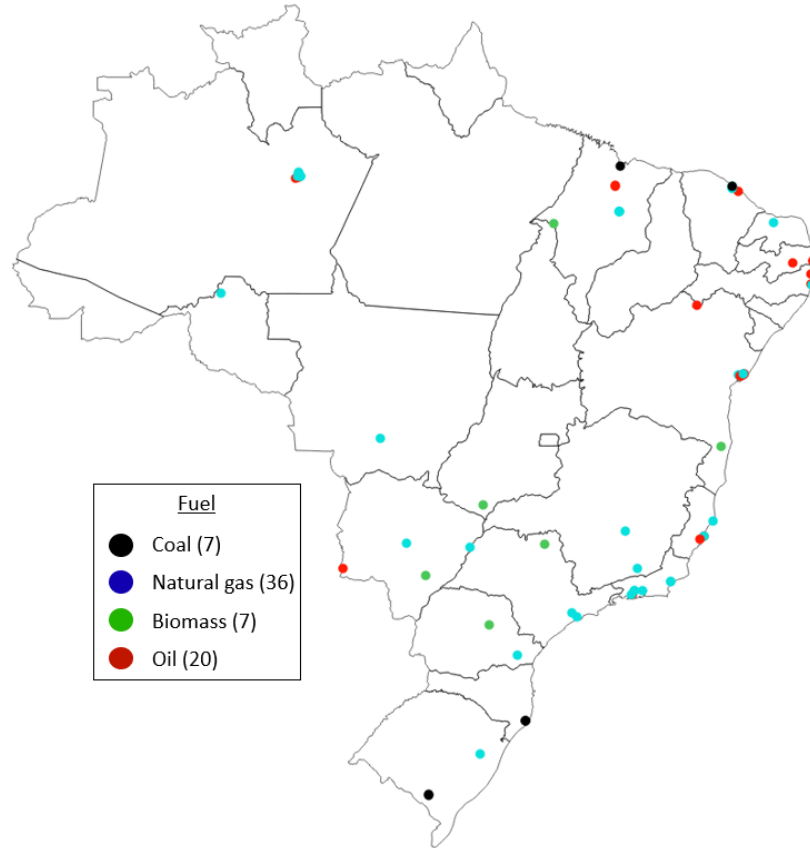
Notes: Location and type of all 70 power plants used in the analysis.

Table 3.3.3 presents the descriptive statistics of the exposure during the pregnancy period. Panel A shows average exposure in all municipalities from the main sample, while Panel B shows the information averaging only among municipalities near a type of power plant.[3.3.13]

[3.3.13]If a municipality is only near a coal power plant, the exposure to natural gas, oil and biomass power plants is always zero. Some municipalities are near more than one type of power plant.

Table 3.3.3 – Exposure to energy generation

Variable	Mean	StdDv	Min	Max	N.mun
Panel A: All municipalities					
Coal power plant	36.17	183.59	0.00	3224.44	500
Natural gas power plant	147.03	397.85	0.00	6312.53	500
Oil power plant	41.72	153.05	-0.00	2993.85	500
Biomass power plant	3.92	36.94	0.00	1046.99	500
Panel B: Only municipalities affected by fuel					
Coal power plant nearby mun	226.04	409.58	0.00	3224.44	80
Natural gas power plant nearby mun	172.57	425.88	0.00	6312.53	426
Oil power plant nearby mun	78.41	202.87	-0.00	2993.85	266
Biomass power plant nearby mun	19.81	81.10	0.00	1046.99	99

Notes: Descriptive statistics of the monthly exposure variable. Panel A shows the average exposure to each type of thermoelectric power plant, considering all the municipalities in the main sample. Panel B also shows the average exposure but considers only cities in the vicinity of each type of power plant.

3.3.4 Other data and sample selection

We use the 2010 Brazilian Census, from the Brazilian Institute of Geography and Statistics (IBGE), to gather the socioeconomic characteristics of the regions exposed to pollution from the power plants and to analyze the population's spatial distribution based on the georeferenced Census Tracts (CTs). To assess the economic activity of each region during the period of pregnancy, we use data from the SIDRA system ("*Sistema IBGE de Recuperação Automática*") from IBGE. They gather information about the municipality's Gross Domestic Product (GDP) at the annual level. We equally divide the annual value through each month of the year, and sum over the months of each pregnancy.

We included in the main sample, only municipalities in the vicinity of at least one of the 70 selected power plants. In this study, we consider a city to be in the vicinity of a power plant, if the city's center lays within 50 km of the facility. To increase precision in selecting the sample, we calculate the center of a municipality as the point with the city's highest population density. To this end, we use information about population distribution from the Oak Ridge National

Laboratory (ORNL).[3.3.14] Their LandScan data shows yearly estimations of the average population at a resolution of approximately one square kilometer for the whole earth. We consider the centroid of the 3 km by 3 km square with the highest population as the center of a municipality.

Large municipalities might have a different dynamic of air pollution, therefore, the cities in the top 5% of the population distribution are excluded, so that the main sample contains 500 municipalities analyzed at the monthly level. Table 3.3.2 presented some characteristics of these municipalities selected to the main sample.

3.4   Methodology

This section presents the methodology used to investigate the effect of an operational power plant on health outcomes. The main explanatory variable is the exposure to power plants presented in the previous section, but there is the possibility of some confounder effects impacting the results. Initially, the weather conditions might influence the fetal health, the dispersion of pollution, and the use of thermoelectric power (DESCHÊNES; GREENSTONE; GURYAN, 2009; EYER; WICHMAN, 2018). Besides that, the economic activity might also pose a challenge for the identification strategy if it relates to both the amount of energy generated and the health outcomes. Fortunately, the regulatory framework of the Brazilian energy sector decouples the regional conditions from the demand for local sources of energy, because it is designed to work as a grid centrally dispatched by the ONS. Therefore, it is not likely that the local economy or weather conditions affect the energy production of nearby power plants. Nevertheless, to guarantee these possibilities are accounted for, we include control variables for the local economic activity and weather conditions. To control for the economic activity, we include the monthly GDP of each municipality estimated from the SIDRA data as discussed in the previous section. The weather controls are the minimum, maximum and average temperatures; the total precipitation; and the average wind speed. Deryugina et

---

[3.3.14]https://landscan.ornl.gov/

al. (2019) show that air pollution can increase when the wind blows from a more polluted region to a cleaner one. Therefore, similar to Rangel e Vogl (2019), the main specification also includes a control for the predominant wind direction in each municipality. All of these control variables will be aggregated to the nine months of pregnancy of babies born at each month (m) and municipality (i). The benchmark specification is presented in Equation 3.4.1.

$$IMR_{im} = \sum_{\forall f} \beta^f E_{im}^f + X_{im}'\delta + \omega_{im} + \phi_{gc} + \lambda_m + \varepsilon_{im} \qquad (3.4.1)$$

IMR_{im} is the infant mortality rate in municipality i and month m. E_{im}^f is the exposure to power plants fueled by f during pregnancy of children born at month m in municipality i. X_{im} includes the weather and GDP controls related to the pregnancy period of babies born at month m in municipality i. X_{im} also controls for the average characteristics of pregnancies of babies born at month m and municipality i. These pregnancy characteristics are the average share of doctor visits before birth, the share of twins birth, the share of regular birth (or share of cesarean), and mother's average quantity of prior pregnancies. Infant mortality in the first year of life might also be affected by exposure to pollution in the months after birth, therefore, we also include in X_{im} the average exposure in the 12 months after birth. ω_{im} controls for the predominant wind direction in municipality i during the pregnancy period of babies born at month m. This variable represents the share of days during the pregnancy period of babies born at month m that the average wind direction blew in each octant of the wind rose at municipality i (so there will be eight variables per municipality, one for each octant). ϕ_{gc} is a calendar-month (c) by municipality block (g) fixed effect that controls for regional seasonality effect. These municipality blocks are defined as municipalities near the same power plant or group of power plants.[3.4.1] λ_m are month fixed effects to capture common time effects, and ε_{im} represents the error term.

[3.4.1]When facilities are too close to each other, it doesn't make sense to assume that municipalities close to one or to another will observe distinct seasonal patterns, so municipalities near power plants within 30 km from one another are considered to be in the same block. There are 37 blocks of municipalities in the the sample, with an average of 13.5 cities per block.

The coefficients of interest in this specification are the β^f, that measure the effect of increasing exposure to f type power plants on the infant mortality rate. An increase in 1 GWh generated by a f-fueled thermoelectric power plant that affects a full municipality will cause an additional β^f deaths by every thousand live births. Observations in the regression are weighted by the average live births in each municipality, and the standard errors are clustered at the level of the municipality block g to account for spatial dependence in the model.

The identification strategy relies on the assumption that the municipality exposure to energy generation is conditionally unrelated to unobserved determinants of birth outcomes. This assumption seems credible because there are two sources of plausibly exogenous variation in the exposure variable. The first is the operation schedule of each power plant defined by the central authority to optimize the system, and the second is the wind direction blowing in each facility.

As discussed in Section 3.3.3, we only consider the large thermoelectric power plants from the ONS data. Nevertheless, smaller facilities generating energy, within or without the SIN, might also pollute the environment with health consequences to local communities. This poses a challenge to the identification strategy if the operation of these small power plants correlates with the generation from the considered power plants and also influences the health of the population. Although this is a relevant threat to the identification, the amount of energy generated by this group of power plants is much smaller than the energy from the included power plants. Besides that, the calendar-month (c) by municipality block (g) fixed effect (ϕ_{gc}) controls for constant characteristics within a region and calendar-month. Therefore, reducing the problem of facilities with cyclical patterns of operation.

3.5   Results

Table 3.5.1 presents the main results of this analysis as well as the gender heterogeneity. In columns 1 through 5, the dependent variables are the infant mortality rate on the first day, first week, first month, first three months, and first year of life. In columns 6 and 7 the analysis focus on female and male deaths.

Results suggest that only coal-fueled power plants led to an increase in the infant mortality rate, which is consistent with this type of facility being one of the largest stationary sources of air pollution. The mean value of the dependent variables, presented in the lower part of the table, shows that around half of the infant mortality rate happens in the first week of life, while the rest is equally spread in the other time frames.

Table 3.5.1 – Infant mortality rate

| | Infant mortality rate (by thousand live births) | | | | | | |
| | 1d | 1w | 1m | 3m | 1y | Female 1y | Male 1y |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.002**	0.0001	0.002***	0.003*	0.003*	0.008***	−0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
Preg. n. gas exp.	0.0002	−0.0002	0.0003	0.0001	0.001	0.001	−0.001
	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	−0.001	0.0002	0.001	−0.0002	−0.001	−0.003	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Preg. bio exp.	−0.004	0.009	0.012	0.004	−0.003	0.006	−0.015
	(0.011)	(0.014)	(0.014)	(0.011)	(0.010)	(0.018)	(0.010)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	13.05
Observations	35,701	35,701	35,701	35,701	35,701	34,970	35,043
Adjusted R^2	0.030	0.032	0.032	0.034	0.031	0.007	−0.002

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for only female, and only male deaths in the first year of life. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Columns 6 and 7 of Table 3.5.1 present the gender heterogeneity. The dependent variables in these specifications are defined as female (male) deaths by a thousand female (male) live births. The estimated coefficients suggest that only girls are affected by the in utero exposure to energy generation. The higher effects observed for female children is robust in different analysis throughout this

work, and, although not typical in this literature, this result finds support in some exercises from the work of Tanaka (2015). He finds larger effects of air pollution on the health of female newborns, but the only statistically significant difference in his exercise appears in the birth weight. We conduct some robustness analysis in the appendices of this work. In Section 3.B, we test the effects of the exposure variable on the infant mortality rate in the first year of life using different model specifications and we test birth outcomes as the dependent variable. In Section 3.C, we use different definitions for the economic activity controls, but the results are similar to those presented in the Table 3.5.1. In Section 3.D, we include controls for the agricultural activity and nearby fire spots, and Section 3.E investigates the exposure in different trimesters of the pregnancy. Results are fairly stable in all of these exercises, but next section argues that fetal selection could be biasing these results toward zero.

3.5.1   Fetal selection

A common concern from the literature analyzing the effect of shocks during pregnancy on health outcomes is the possibility of fetal selection, discussed in Section 3.1. To investigate the extent to which selection bias affects the main results, we analyze if exposure to energy generation during pregnancy influences the incidence of fetal deaths in the surrounding municipalities (similar to Rocha e Soares (2015)). Columns 1 through 3 analyze slightly different definitions of fetal death. Column 1 estimates the effect of exposure to power plants on the number of fetal deaths (in level). The dependent variable in columns 2 is the fetal death rate calculated by dividing fetal deaths by the number of potential births, defined as the sum of the observed live births with fetal deaths normalized by a thousand. Column 3 uses as dependent variable a measure of potential infant mortality rate that is calculated by dividing total deaths (fetal death plus infant deaths) by the potential births (fetal death plus observed live birth). Columns 4 and 5 repeat this potential infant death specification from column 3, but focusing on the mortality of, respectively, girls only and boys only.

Table 3.5.2 – Fetal death

	Dependent variable:				
	F.D. lvl	F.D. rate	F.D. + 1y rate	Female: F.D. + 1y rate	Male: F.D. + 1y rate
	(1)	(2)	(3)	(4)	(5)
Preg. coal exp.	0.0005**	0.005*	0.008***	0.015***	−0.003
	(0.0002)	(0.003)	(0.003)	(0.003)	(0.006)
Preg. n. gas exp.	−0.0001	−0.0003	0.0003	0.001	0.001
	(0.0002)	(0.001)	(0.001)	(0.001)	(0.002)
Preg. oil exp.	0.00002	0.0003	−0.001	−0.002	0.002
	(0.0003)	(0.002)	(0.002)	(0.002)	(0.002)
Preg. biomass exp.	0.001	−0.003	−0.006	0.011	−0.032**
	(0.0005)	(0.008)	(0.005)	(0.020)	(0.013)
Mean dep. var	2.67	10.49	22.48	19.19	21.86
Observations	35,701	35,701	35,701	34,976	35,049
Adjusted R^2	0.581	0.037	0.058	0.034	0.034

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. Column 1 shows the results for the number of fetal deaths. Column 2 presents the results for the fetal death divided by the potential live births in each month and municipality, normalized by a thousand. Column 3 uses as dependent variable total deaths (fetal deaths + infant mortality) divided by potential births (fetal death + live births), normalized by a thousand. Columns 4 and 5 repeats specification from column 3, but divides the sample between girls only and boys only deaths. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

These results suggest that fetal selection is an important factor affecting the main estimations. Exposure to coal power plant during pregnancy seems to be inducing an early termination of pregnancies, especially for female fetuses (column 4). Therefore, the main results presented in Table 3.5.1 are likely underestimating the real impact of thermoelectric power on infant mortality rate (results considering fetal death are around three times larger than the benchmark specification). It's worth noting that this measure of fetal deaths is also an underestimation of the real life loss since fetal deaths are only observed after a certain gestational age. This table also has an unexpected result for male children exposed to biomass energy during pregnancy. It suggests that the higher the exposure to biomass power plants, the lower the males' potential infant mortality rate.

3.5.2 Infant mortality rate by cause of death

In the previous section, we acknowledged that fetal selection is likely biasing the observed results. This section breaks the main exercise of the infant mortality rate in the first year of life (Column 5 of Table 3.5.1) into different registered causes of death. Column 1 repeats the main analysis showing results for all deaths of children under one year of age. Column 2 shows the results for deaths caused by diseases of the respiratory system. Column 3 shows the estimations for conditions that originated in the perinatal period, while column 4 presents a subgroup of these diseases that relates to respiratory and cardiovascular issues. Column 5 shows deaths related to congenital malformation, and column 6 is a subgroup containing respiratory and cardiovascular conditions. Column 7 is a robustness check, where we test the effect of power plants on deaths unrelated to air pollution. It shows deaths related to external causes such as accidents, assault, complications from medical and surgical care.

Table 3.5.3 – Infant mortality by cause

	Dependent variable:						
	All dths	Resp.	Per	Per R-C	MF	MF C	Ext
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.003*	−0.0001	0.001	0.001	0.003*	0.001	−0.0003
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.0003)
Preg. n. gas exp.	0.001	0.0004**	0.001	0.0003	−0.001	−0.0004	0.0001
	(0.001)	(0.0002)	(0.001)	(0.001)	(0.0004)	(0.0003)	(0.0002)
Preg. oil exp.	−0.001	0.00003	0.002**	0.002**	−0.002**	−0.001*	−0.0003
	(0.001)	(0.0003)	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0003)
Preg. biomass exp.	−0.003	−0.002	0.008	−0.004	0.008*	0.011***	0.0003
	(0.010)	(0.001)	(0.010)	(0.005)	(0.005)	(0.002)	(0.0004)
Mean dep. var	12.12	0.63	6.96	2.63	2.73	1.29	0.4
Observations	35,701	35,701	35,701	35,701	35,701	35,701	35,701
Adjusted R^2	0.031	0.013	0.035	0.023	−0.006	−0.012	0.012

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. Column 2 considers deaths related to diseases of the Respiratory system (ICD-10 codes that start with the letter "J"). Column 3 contains deaths related to certain conditions originating in the Perinatal period (ICD-10 codes that start with the letter "P"). Column 4 shows deaths related to Respiratory and Cardiovascular disorders specific to the perinatal period (ICD-10 codes "P20-29"). Column 5 shows deaths related to congenital malformations (MF), deformations and chromosomal abnormalities (ICD-10 codes that start with the letter "Q"). Column 6 shows deaths related to congenital malformations of the Respiratory and Circulatory system (ICD-10 codes "Q20-28" and "Q30-34"). Column 7 shows results of deaths related to external causes and accidents (ICD-10 codes that start with "V", "W", "X", and "Y"). Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

The bottom row of Table 3.5.3 shows the mean value of the dependent variables. It suggests that around 5% of the mortality rate in the first year of life are due to respiratory diseases (0.63 out of 12.12 deaths), half of it comes from conditions that originated in the perinatal period (6.96 out of 12.12 deaths per live births), and another 20% comes from congenital malformation (2.73 out of 12.12). The robustness check (column 7) analyzes deaths unrelated to air pollution. This type of death represents only a small share of the total observed infant mortality, but it does not respond to any sort of nearby energy generation. The observed effect from exposure to coal energy generation comes mainly from deaths due to congenital malformation, deformations, and chromosomal abnormalities (column

5), therefore, in Table 3.5.4, we investigate heterogeneous effects of this disease in different periods of life and gender.

Table 3.5.4 – Age and gender heterogeneity for congenital malformation diseases

	Mortality rate by malformation				
	1d	1w	1m	3m	1y
	(1)	(2)	(3)	(4)	(5)
Male children					
Preg. coal exp.	0.0001	0.0004	−0.002	−0.002	−0.001
	(0.001)	(0.001)	(0.003)	(0.004)	(0.004)
Preg. n. gas exp.	−0.0001	−0.0003	−0.001*	−0.001*	−0.001
	(0.0002)	(0.0004)	(0.0004)	(0.001)	(0.001)
Preg. oil exp.	0.0003	−0.0004	−0.001	−0.001*	−0.002*
	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	0.001	0.014***	0.011**	0.015***	0.012*
	(0.003)	(0.005)	(0.005)	(0.005)	(0.006)
Mean dep. var	0.59	1.31	1.81	2.4	2.81
Observations	35,043	35,043	35,043	35,043	35,043
Adjusted R^2	0.029	0.014	0.008	−0.005	−0.019
Female children					
Preg. coal exp.	0.002***	0.001**	0.002***	0.002**	0.004***
	(0.001)	(0.0005)	(0.0004)	(0.001)	(0.001)
Preg. n. gas exp.	0.0001	0.0002	0.0001	−0.0001	−0.0004
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Preg. oil exp.	−0.001**	−0.0002	0.0001	−0.001	−0.001
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	0.003	0.004	−0.001	0.001	0.005
	(0.003)	(0.005)	(0.007)	(0.006)	(0.008)
Mean dep. var	0.57	1.15	1.58	2.17	2.62
Observations	34,970	34,970	34,970	34,970	34,970
Adjusted R^2	0.013	−0.001	−0.002	−0.005	−0.008

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable is infant death by congenital malformations (MF), deformations and chromosomal abnormalities (ICD-10 code starts with the letter "Q"). Each columns analyzes a specific period of death, and each panel shows the results for a specific gender. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

The results observed in the mortality from congenital malformation are mixed. While male children seem to be more affected by exposure to biomass power plants, female kids are more affected by facilities fueled by coal. The effects observed in this table are the main drivers of the adverse impact of coal-fueled power plants on female children's health. Appendix Section 3.G further explores the mortality from each disease by gender and period of life.

3.5.3   Distance from power plant

Table 3.5.5 analyzes the heterogeneous effects of each type of power plant according to the distance between a facility and the affected population. A caveat from this exercise is that the infant mortality rates used as dependent variables are only measured at the municipality level. Therefore, we will slightly change the explanatory (but not the dependent) variables to assess if exposure to a power plant in a given distance range has more or less influence on the results. We consider the share of the population exposed to a power plant within 20 km, between 20 and 40 km, and above 40 km. This variable is built like in Equation 3.3.1, but interacting it with another categorical variable saying if each Census Tract (s) is within each of the distance groups of the power plant (j). We present this new explanatory variable in Equation 3.5.1.

$$E_{id}^{fr} = \sum_{\forall j \in i} \sum_{\forall j \in i} \sum_{\forall r \in \{0-20, 20-40, 40+\}} G_{jd}^{f} * W_{sjd}^{i} * p_{s,i} * D_{sj}^{r} \qquad (3.5.1)$$

This equation is similar to the previously presented model (Equation 3.3.1), but includes an interaction with an indicative variable (D_{sj}^{r}) that equals 1 if the distance between Census Tract s and power plant j is within the distance group r.

The results suggest that the increase in the infant mortality rate from exposure to coal power plants are driven by the population living near this type of facility. In most of the time frames analyzed, the coefficient to coal power plant within 20 km is larger than the other distance groups and is the only statistically significant coefficient. Female children living near a coal power plant also seem to

be largely affected by the energy generation, but male kids appear to have a large and unexpected signed coefficient. Columns 7 and 9 also present the exercise using the potential death outcomes (which considers fetal death) as dependent variables. Fetal deaths also seem to be affecting the results for female children living near the power plants. Besides coal power plants, biomass facilities show some mixed results depending on the distance group analyzed, but the aggregated effects, presented in Table 3.5.1, are statistically equal to zero.

Table 3.5.5 – Infant mortality rate by distance to the emitting source

| | *Dependent variable:* | | | | | | | | |
| | 1d | 1w | 1m | 3m | 1y | Fem 1y | Fem. ftl. dth. | Male 1y | Male ftl. dth |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal exp.: up to 20 km	0.004***	0.0003	0.006*	0.008***	0.009***	0.032***	0.042***	−0.016***	−0.005
	(0.001)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.010)
20 to 40 km	−0.001	−0.003	0.0003	−0.003	−0.001	−0.005	−0.0004	0.002	−0.003
	(0.002)	(0.002)	(0.005)	(0.004)	(0.003)	(0.003)	(0.009)	(0.007)	(0.009)
above 40 km	0.004***	0.002	0.003	0.005	0.005	0.011**	0.019***	0.0005	−0.003
	(0.001)	(0.002)	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.006)	(0.009)
N.gas exp.: up to 20 km	0.001*	−0.0001	−0.0001	−0.001	−0.00002	−0.001	−0.0005	0.0001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)	(0.003)
20 to 40 km	−0.00002	−0.0003	0.001	0.001	0.002**	0.002*	0.002	0.002	0.003
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
above 40 km	−0.0001	0.0001	−0.0002	−0.001	−0.002	0.003	0.001	−0.008**	−0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Oil exp.: up to 20 km	0.0002	0.003**	0.002	−0.0004	−0.001	0.003	−0.001	−0.0005	0.004
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.003)
20 to 40 km	−0.003	−0.002	0.001	0.001	0.001	−0.005	−0.004	0.006	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.006)
above 40 km	0.001	−0.0002	−0.001	−0.002	−0.004	−0.006	0.001	0.007	0.001
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.006)	(0.009)	(0.009)
Bio exp.: up to 20 km	−0.025**	−0.023	−0.030**	−0.014	−0.017	0.001	−0.033	−0.032	−0.125***
	(0.009)	(0.022)	(0.011)	(0.013)	(0.013)	(0.016)	(0.063)	(0.046)	(0.032)
20 to 40 km	0.006	0.026***	0.029***	0.020***	0.011***	0.019	0.025***	−0.001	−0.023
	(0.007)	(0.003)	(0.003)	(0.002)	(0.003)	(0.011)	(0.009)	(0.004)	(0.015)
above 40 km	−0.039***	−0.052***	−0.047***	−0.059***	−0.056***	−0.046***	−0.036***	−0.066***	−0.033
	(0.007)	(0.006)	(0.008)	(0.010)	(0.012)	(0.009)	(0.013)	(0.017)	(0.033)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	19.19	13.05	21.86
Observations	35,701	35,701	35,701	35,701	35,701	34,970	34,976	35,043	35,049
Adjusted R^2	0.030	0.032	0.032	0.034	0.031	0.007	0.034	−0.002	0.034

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. All dependent variables are normalized to a thousand live-births. The dependent variable in Columns 1, 2, 3, 4, and 5 show estimations for the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for female death and the potential female infant mortality, that considers the fetal death. Columns 8 and 9 show results for male death and for the potential male infant mortality. Main explanatory variables are the exposure in a given distance range to each type of power plant during the pregnancy period.

98

3.5.4   Hospitalization

In this section, we discuss the effects of using thermoelectric power plants on hospitalization rates for the population in different age groups.[3.5.1] The hospitalization data from Datasus indicates the zip code of residence of each hospitalized patient, therefore, in this analysis, we will use a geographical unit of observation smaller than a municipality. The "*CEP Aberto*" project[3.5.2] collects, in collaboration with the community, information from Brazilian zip codes such as the state, municipality, street name, altitude, latitude, and longitude. We used this data to allocate hospitalizations from the zip code of residence of a patient to the Census Tract (CT) that contains it. We used the same municipality sample selection as in the previous section, and, from all hospitalization in this sample, around 10% could not be found in the "*CEP Aberto*" data and were dropped from the analysis. Figure 3.5.1 shows the hospitalization rate by respiratory diseases from the main sample selection compared to the country as a whole.[3.5.3] The rates are in the same order of magnitude and follow a similar trend.

[3.5.1]In the previous exercises, we investigate the consequences of exposure during pregnancy, while in this section, we analyze if a month with higher exposure leads to higher hospitalization in the same period.

[3.5.2]https://cepaberto.com/

[3.5.3]We considered all hospitalizations containing the information about the zip code of residence.

Figure 3.5.1 – Monthly hospitalization rate by respiratory diseases in different regions
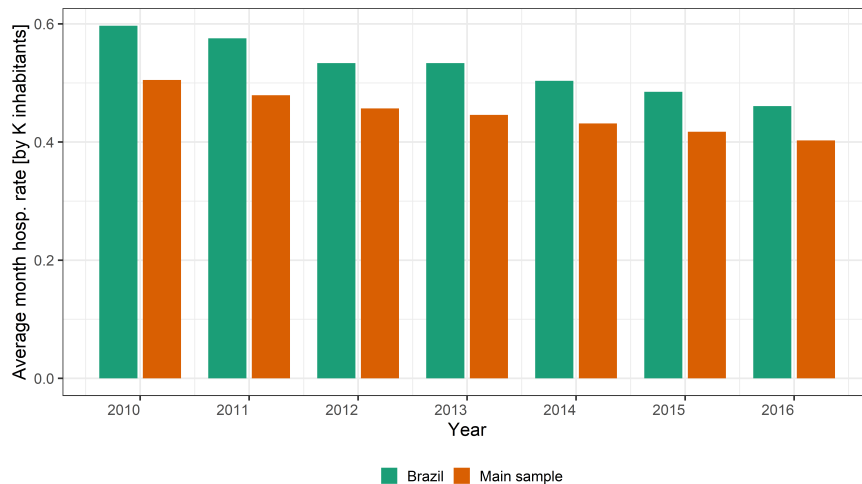
Because in this analysis the unit of observation is a CT, we will slightly change the exposure variable to work with this more refined geographical aggregation. Instead of calculating the share of the population exposed to a power plant, we will consider that everyone from the CT is exposed to a power plant if its centroid is downwind of the facility. This variable is presented in Equation 3.5.2.

$$E_{sm}^{f} = \sum_{\forall j \in i} G_{jm}^{f} * W_{jm}^{s} \tag{3.5.2}$$

where E_{sm}^{f} is the exposure of Census Tract s at month m to energy generated by fuel f. G_{jm}^{f} is the energy generated by power plant j, fueled by f, at month m. W_{jm}^{s} is an indicative variable equals 1 if the Census Tract s is within a range of 45° of the average wind direction at month m. Like in the previous definition of the exposure variable, we sum over all power plants within 50 km of the center of municipality i that contains the Census Tract s. Table 3.5.6 presents the main characteristics of the sample of Census Tracts and also the exposure to each type of energy.

Table 3.5.6 – Monthly exposure to energy generation

Variable	Mean	StdDv	Min	Max
Census Tract characteristics				
Total population	646.26	367.43	1.00	4020.00
Pop. age: under 1	9.51	7.19	0.00	76.00
Pop. age: under 5	58.40	40.06	0.00	471.00
Pop. age: 6 to 64	555.46	310.60	1.00	3546.00
Pop. age: above 65	42.48	28.23	0.00	318.00
Share pop. below 1 Min Wage	0.63	0.24	0.00	1.00
Exposure to energy generation				
Coal power plant	3.29	20.51	0.00	354.36
Natural gas power plant	22.69	48.53	0.00	965.50
Oil power plant	4.97	17.04	-0.00	371.32
Biomass power plant	0.12	2.22	0.00	131.03

Notes: Descriptive statistics of Census Tract characteristics and monthly exposure to a nearby operating power plant. The main sample contains 43,706 Census Tracts. The variable "Share pop. below 1 Min Wage" is the average share of the population from the Census Tracts with a wage per capita below the minimum wage.

In this section, we also slightly change the main specification to consider that we are working with smaller observation units. CTs that belong to the same municipalities might experience similar shocks, therefore, we include controls and fixed effects to capture common monthly characteristics of all CTs within the same municipality. The regressions are weighted by the population of each Census Tract within an age group, and the standard errors are clustered at the municipality level. The main specification used in this exercise is given by Equation 3.5.3.

$$H_{sm} = \sum_{\forall f} \beta^f E_{sm}^f + X'_{sm}\delta + \sigma_s + \omega_{im} + \phi_{gc} + \alpha_{im} + \varepsilon_{sm} \qquad (3.5.3)$$

where H_{sm} is the hospitalization rate at month m and CT s. E_{sm}^f is the exposition to energy defined by Equation 3.5.2. X_{sm} includes weather control variables for CT s at month m. The weather control variables are total precipitation, mean wind speed, mean, maximum, and minimum temperature during the month m.[3.5.4]

[3.5.4]To calculate weather conditions for each Census Tract we used the same algorithm as

σ_s is a Census Tract fixed effect. ω_{im} is a variable that controls for predominant wind direction in each municipality, which is similar to the previous specification. For computational reasons, instead of considering the share of days the wind in a municipality blew in each octant of the wind rose, we divide the wind directions into quadrants, therefore, we have half the number of variables. ϕ_{gc} is a block of municipalities by calendar-month fixed effect to capture regional seasonality. α_{im} is a municipality by month fixed effect to capture common trends in CTs from the same city, and ε_{sm} is the error term.

Table 3.5.7 shows the hospitalization by all respiratory diseases in different age groups. Most of the coefficients are indistinguishable from zero, except for the borderline significant results with unexpected signs for the exposure to oil power plants in patients under one year of age. Appendix Table 3.I.4 suggests that hospitalization by pneumonia drives this unexpected result.

Table 3.5.7 – Hospitalization by all respiratory diseases in different age groups

	All ages (1)	Age 0-1 (2)	Age 0-1: male (3)	Age 0-1: fem. (4)	Age 0-5 (5)	Age 6-64 (6)	Age 65+ (7)
Coal expo	-0.00028	-0.00692	-0.00442	-0.00251	-0.00163	-0.00012	-0.00364
	(0.00026)	(0.01093)	(0.00694)	(0.00469)	(0.00196)	(0.00018)	(0.00299)
Ngas expo	-0.00016	-0.00291	-0.00027	-0.00264	-0.00087	-0.00010	-0.00021
	(0.00023)	(0.00480)	(0.00303)	(0.00205)	(0.00097)	(0.00021)	(0.00046)
Oil expo	-0.00066*	-0.01304*	-0.00884**	-0.00420	-0.00548**	-0.00018	-0.00072
	(0.00036)	(0.00686)	(0.00405)	(0.00319)	(0.00240)	(0.00014)	(0.00087)
Bio expo	0.00149	0.01791	0.01250	0.00541	0.00791	0.00073*	0.00150
	(0.00129)	(0.01804)	(0.01030)	(0.00885)	(0.00897)	(0.00044)	(0.00329)
Mean dep.var	0.37	6.44	3.77	2.66	2.03	0.18	1.07
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.77	0.55	0.52	0.45	0.61	0.71	0.60

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the number of inhabitants in each Census Tract in each age group. Standard-errors clustered at municipality level. Each column shows the result for hospitalization by all types of respiratory disease in a specific age group. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

---

presented in Section 3.3.2.

3.6   Discussion

Brazil is a large country that relies heavily on hydroelectric power in its energy mix, but in recent years the use of thermoelectric power plants has increased from 5% in the 2000s to something around 20% by 2017. Adverse weather shocks that compromised the capacity of water reservoirs supplying hydroelectric facilities and investments in other types of energy generation - mainly thermoelectric and wind facilities - help explain this shift away from hydroelectric power. In this study, we investigate if the use of thermoelectric power plants affects health outcomes of nearby communities.

We conduct two exercises. The first one analyzing the infant mortality rate at the municipality level, and the second one evaluating hospitalization at the Census Tract level. We built an index variable to measures the exposure of each geographical unity (municipality or Census Tract) to nearby operational power plants. The results suggest that exposure to coal power plants during pregnancy can lead to an increase in the infant mortality rate, but there is no clear evidence of increasing hospital admissions in months with higher use of thermoelectric power.

Only coal-fueled power plants seem to lead to an increase in the infant mortality rate, but the coefficients' magnitude is low. In the benchmark specification, every additional gigawatt-hour of energy affecting a whole municipality increased the mortality in 0.003 deaths by a thousand live births in a month, from an average mortality rate of 12. In our results, this negative effect of coal power plants is highly concentrated in female children.

The low coefficients observed in the results should be interpreted with caution because there is evidence that fetal selection plays an important role in biasing down the estimations. The use of thermoelectric power plants might lead to an early termination of pregnancies, which must also be treated as an adverse health consequence of this type of facility. Besides that, these negative consequences are not homogeneously spread through space, and the results suggest that the population living closer to power plants might be more affected by their pollution.

Brazil is increasingly relying on thermoelectric power to supply its energy

demand, but most of the new investments focus on natural gas facilities. Although this shift toward thermoelectric power should raise concerns regarding the increase of emissions, it is reassuring that the country does not rely entirely on coal power plants since these facilities are more harmful to the environment and the nearby population's health. As further developments of this research, it would be important to expand the sample of thermoelectric power plants to contemplate more facilities operating in the country. Besides that, the length of exposure to an operational power plant might influence the health of the affected population, therefore, it would be interesting to include this variable in the analysis.

Appendix

3.A Descriptive statistics

Figure 3.A.1 shows the evolution of infant mortality rates for children up to one day, one week, one month, three months, and one year of life from 2010 to 2016. There is a slight decrease in these rates, but this variation is small.

Figure 3.A.1 – Evolution of the infant mortality rate
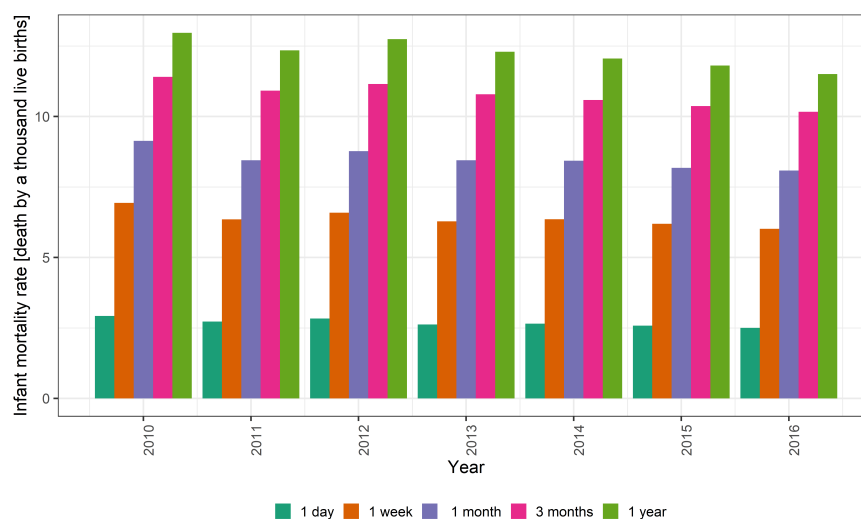
Notes: This figure shows the evolution of the infant mortality rate in different periods of the first year of life including the 500 municipalities from the main sample.

Figure 3.A.2 shows the distribution of the population in the vicinity of power plants according to the distance between the Census Tract of residence and the facility.

Figure 3.A.2 – Share of the population according to distance to a power plant
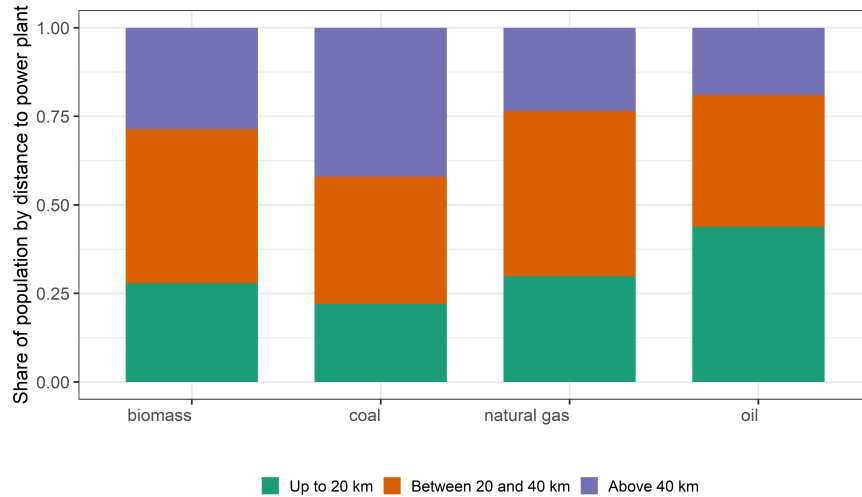
Notes: Figure showing the distribution of population in the vicinity of power plants including the 500 municipalities from the main sample.

Figure 3.A.3 presents the share of days (after the first operational day) that each power plant is operational by its maximum observed daily generation capacity (on the x axis). There is a large variation in the share of operational days. Larger facilities (to the right of the Figure) are used more often, but a large share of the power plants from the sample is frequently turned on and off.

Figure 3.A.3 – Power plants' operational days
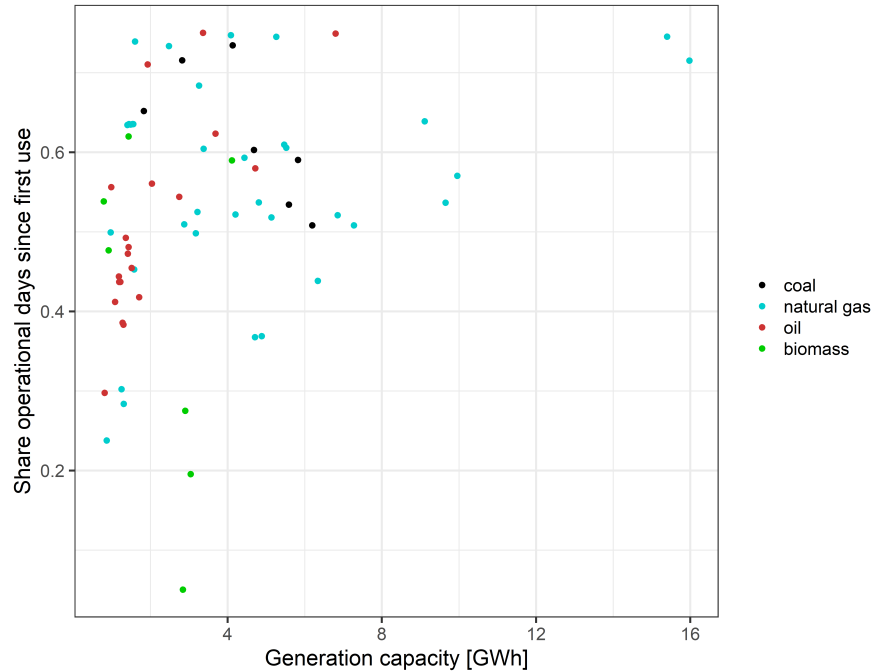
Notes: This figure shows the share of days each power plant operated (generation larger than zero) since the first operational day according to the generating capacity. The capacity axis (x) shows the maximum one-day observed generation in the sample.

3.B Alternative specification

Table 3.B.1 shows different model specifications for the mortality rate in the first year of life to compare them to the benchmark specification presented in column 9. Once we include the municipality fixed effect (in column 5), the coefficients for the exposure to coal and to natural gas become positive and significant. The coefficient for coal exposure remains almost unchanged in all the other specifications, but the coefficient to exposure to natural gas losses significance after including controls for predominant wind direction (Column 7).

In columns 1 through 10, we use the main sample in which we exclude 27 municipalities (5%) with the largest populations. In column 11, we add these municipalities to the sample to analyze if this exclusion is affecting the results, but the estimated coefficient for exposure to coal remains similar.

Table 3.B.1 – Alternative specification

	Dependent variable:										
	Infant mortality (death by a thousand live births)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Preg. coal exp.	−0.001***	−0.001	−0.001	−0.001	0.002*	0.002*	0.003	0.003*	0.003*	0.004*	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Preg. n. gas exp.	0.0003	0.0004	0.0003	0.0004	0.001*	0.001*	0.001	0.001	0.001	0.0004	−0.0004
	(0.001)	(0.001)	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	0.00002	−0.002	−0.002	−0.002	0.001	0.001	−0.002**	−0.001	−0.001	−0.001	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	−0.003	−0.008	−0.006*	−0.006*	−0.003	−0.004	0.005	−0.003	−0.003	−0.001	−0.001
	(0.005)	(0.005)	(0.003)	(0.003)	(0.005)	(0.005)	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)
Mean dep. var	12.12										
Weather controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
Birth controls			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE					Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mun. block x cal mth FE						Yes	Yes	Yes	Yes	Yes	Yes
Wind dir x city							Yes	Yes	Yes	Yes	Yes
1st year expos								Yes	Yes	Yes	Yes
GDP control									Yes	Yes	Yes
Weather bins										Yes	
Add big cities											Yes
Observations	35,701	35,701	35,701	35,701	35,701	35,701	35,701	35,701	35,701	35,701	37,645
Adjusted R^2	0.0001	0.003	0.015	0.017	0.036	0.035	0.031	0.031	0.031	0.031	0.056

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard-errors clustered at the municipality block level. All columns show estimation for the same dependent variable, but controls and fixed effects are added in each column according to bottom rows. Column 9 is the benchmark specification, with birth variables and weather controls, regional specific seasonality, month fixed effect, and with controls for municipality specific wind direction. Column 10 repeats the benchmark specification, but instead of using average weather variables as controls, it divides the weather variables in frequency within bins. Column 11 returns to the benchmark specification but uses the full sample, without excluding the largest municipalities. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.B.2 shows the estimation of the effect of the exposure on different birth outcomes. Results for coal exposure on the share of Low Birth Weight (column 2) and of Pre-Term (column 4) are statistically significant, but the coefficients have an unexpected sign and are small. These coefficients pointing to the direction opposite to what would be expected could reflect the fetal selection affecting birth outcomes. If this is the case, because the weaker fetuses died before birth, only the stronger individuals ended up being born.

Table 3.B.2 – Birth outcomes

	Dependent variable:						
	Weight	LBW	VLBW	PT	VPT	APGAR1	APGAR5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.008	−0.00001***	−0.00000	−0.00002**	0.00000	0.00002	0.00003
	(0.018)	(0.00000)	(0.00000)	(0.00001)	(0.00001)	(0.00005)	(0.0001)
Preg. n. gas exp.	0.002	0.00000	0.00000	−0.00000	0.00000	−0.00003	−0.00005
	(0.004)	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.00005)	(0.00004)
Preg. oil exp.	0.005	−0.00000	−0.00000	0.00002	−0.00000*	−0.0002	−0.0001
	(0.010)	(0.00000)	(0.00000)	(0.00001)	(0.00000)	(0.0001)	(0.0001)
Preg. biomass exp.	0.025	−0.00000	0.00001	0.00001	0.00003**	−0.001***	−0.0004**
	(0.081)	(0.00002)	(0.00001)	(0.00004)	(0.00001)	(0.0002)	(0.0001)
Mean dep. var	3181.36	0.09	0.01	0.11	0.02	8.37	9.34
Observations	35,701	35,701	35,701	35,701	35,701	35,665	35,665
Adjusted R^2	0.389	0.192	0.059	0.240	0.057	0.425	0.484

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard-errors clustered at municipalities near the same power plant or group of nearby power plants. Column 1 shows results for the average birth weight. Columns 2 and 3 presents the share of low birth weight (under 2500 g) and very low birth weight (under 1500g). Columns 4 and 5 are the results for pre-term (under 36 weeks), and very pre-term (under 32 weeks). Columns 6 and 7 show results for APGAR 1 and APGAR5. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

3.C   Alternative controls for the economic activity

In this section, we use alternative methods to further investigate if the economic activity might be interfering with the results. In Table 3.C.1, we replace the GDP control variable for an alternative variable that captures the economic activity. We use public data gathered by the Brazilian Central Bank that presents monthly

information of the banking activities in each municipality. It contains information about the municipalities' total cash availability, banking deposits, securities, real estate bonds, credit operations, among others. We use the total cash availability and credit operations within a month as a proxy for the monthly economic activity in a given municipality. However, a caveat is that only municipalities with banks appear in the data. We used these banking variables as controls in the main regressions and considered them as zero in the municipalities not showing in the data. The results, presented in Table 3.C.1, are equal to those observed in the main specification, at least up to the decimal precision presented here.[3.C.1]

Table 3.C.1 – Infant mortality rate controlling for banking variables

	Infant mortality rate (by thousand live births)						
	1d	1w	1m	3m	1y	Female 1y	Male 1y
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.002**	0.0002	0.002***	0.003*	0.003*	0.008***	−0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
Preg. n. gas exp.	0.0002	−0.0002	0.0003	0.0001	0.001	0.001	−0.001
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	−0.001	0.0003	0.001	−0.0001	−0.001	−0.003	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Preg. bio exp.	−0.003	0.009	0.012	0.004	−0.002	0.006	−0.014
	(0.011)	(0.014)	(0.013)	(0.011)	(0.010)	(0.018)	(0.009)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	13.05
Observations	35,701	35,701	35,701	35,701	35,701	34,970	35,043
Adjusted R^2	0.030	0.032	0.031	0.034	0.031	0.007	−0.002

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for female only, and male only deaths in the first year of life. Main explanatory variables are the exposure to each type of power plant during the pregnancy period. This specification replaces the GDP control for banking variables.

We conduct yet another exercise to investigate if the economic activity affects

---

[3.C.1]Differences between the results appear when we increase the decimal places.

the observed results. To disentangle the energy production in each power plant from the local economic activity and weather, we estimate the energy generation from a thermoelectric power plant in a given month based on the water available to all hydroelectric power plants of the system on the previous month. We use monthly data from the ONS on the levels of water available to all hydroelectric power plants of the country to find an estimated energy generation in each thermoelectric facility. After that, we repeat the main exercise of this study, replacing the observed energy generation from Equation 3.3.1 by the estimated values. Formally, we pool the observations from all thermoelectric power plants and estimate Equation 3.C.1 that follows.

$$TG_{jm} = L_{m-1} + \theta_j + L_{m-1} * \theta_j + \varepsilon_{jm} \qquad (3.C.1)$$

where TG_{jm} is the total energy generation from power plant j at month m. L_{m-1} is the level of available water in all hydroelectric reservoirs in the month before m, θ_j is a fixed effect for power plant j, and ε_{jm} is the error term. Some power plants only started to operate in the middle of the period we are analyzing.[3.C.2] In these situations, the generation was defined as a missing value. Figure 3.C.1, compares the sum of the total observed energy generation from all thermoelectric power plants in the sample, with the fitted values estimated by Equation 3.C.1.

[3.C.2]As pointed by Figure 3.2.2.

Figure 3.C.1 – Comparing the observed monthly energy generation with the fitted value

Notes: This figure compares the observed total energy generation from the power plants in the sample to the total generation estimated based on the level of the water reservoirs in the last period.

We have the estimated monthly energy generated by each power plant of the sample, but we still need to find the daily values to use in Equation 3.3.1. We split the estimated monthly values between the days following the same daily generation schedule from the observed data, as shown in Equation 3.C.2.

$$\widehat{G_{jd}} = \widehat{TG_{jm}} * \frac{G_{jd}}{TG_{jm}} \qquad (3.C.2)$$

$\widehat{TG_{jm}}$ is the total estimated energy generation from power plant j at month m. TG_{jm} is the observed value for the energy generation of power plant j at month m. G_{jd} is the observed daily generation from power plant j at day d, and $\widehat{G_{jd}}$ is the estimated daily generation that will be used in Equation 3.3.1 to build the new data set.

Table 3.C.2 shows the estimations of the main specification using the new dataset. The results are similar to those observed in the main exercise. The coefficients of the impact of exposure to coal are slightly smaller and less significant,

but they still point in the same direction. The effect of exposure to natural gas power plants in the first three months of life becomes negative and significant, and there is no relevant difference for the exposure to oil and biomass power plants.

Table 3.C.2 – Infant mortality rate using the estimated energy generation to calculate the exposure variable

| | Infant mortality rate (by thousand live births) | | | | | | |
| | 1d | 1w | 1m | 3m | 1y | Female 1y | Male 1y |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.0001	0.001	0.003**	0.002*	0.002	0.006***	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Preg. n. gas exp.	−0.0004	−0.002***	−0.002**	−0.001**	−0.001	−0.00002	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Preg. oil exp.	0.0001	0.001	0.003*	0.001	0.001	−0.003	0.016
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.011)
Preg. bio exp.	−0.003	0.019	0.019	0.009	0.007	−0.008	0.022
	(0.008)	(0.014)	(0.015)	(0.014)	(0.014)	(0.014)	(0.017)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	13.05
Observations	35,701	35,701	35,701	35,701	35,701	34,970	35,043
Adjusted R^2	0.030	0.032	0.032	0.034	0.031	0.007	−0.002

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for only female, and only male deaths in the first year of life. Main explanatory variables are the exposure to each type of power plant during the pregnancy period estimated by the level of hydroelectric water reservoirs in the preceding period.

3.D Controlling for the agricultural activity and the incidence of fire spots during pregnancy

In the analysis so far, we used the GDP of each municipality to control for the economic activity, but, in this section, we slightly change the main specification to investigate if the agricultural economic activity influences the results. In Table 3.D.1, we use information from SIDRA about a municipalities' agricultural GDP to

build the control variables for the economic activity. We replace the GDP variable for this agricultural GDP, and the results remain almost unchanged.

Table 3.D.1 – Infant mortality rate controlling for agricultural GDP

	Infant mortality rate (by thousand live births)						
	1d	1w	1m	3m	1y	Female 1y	Male 1y
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.002**	0.0002	0.002***	0.003**	0.003**	0.008***	−0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
Preg. n. gas exp.	0.0001	−0.0003	0.0003	0.0001	0.0005	0.001	−0.001
	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	−0.001	0.0002	0.001	−0.0002	−0.001	−0.003	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Preg. bio exp.	−0.003	0.009	0.012	0.004	−0.003	0.006	−0.015
	(0.011)	(0.014)	(0.013)	(0.011)	(0.011)	(0.018)	(0.010)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	13.05
Observations	35,701	35,701	35,701	35,701	35,701	34,970	35,043
Adjusted R^2	0.030	0.032	0.032	0.034	0.031	0.007	−0.002

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for only female, and only male deaths in the first year of life. Main explanatory variables are the exposure to each type of power plant during the pregnancy period. In this specification, the GDP control variable is replaced by the agricultural GDP in each municipality.

In Table 3.D.2, we investigate if the occurrence of fire spots in the vicinity of a municipality during the pregnancy period of a newborn changes the relationship between the use of power plants and the mortality rates. We use data from the program *"Queimadas"* of the Brazilian National Institute for Space Research (INPE).[3.D.1] We downloaded shapefiles with information on daily fire spots and counted the number of spots within 50 km from the center of each municipality for the whole pregnancy period of babies born at a given month. The results are indifferent for the exposure to fossil fuels, but there is a slight difference in the exposure to biomass power plant. The coefficient showing the impact of the

---
[3.D.1]http://queimadas.dgi.inpe.br/queimadas/portal.

exposure to biomass on the mortality of male children, that was negative but indistinguishable from zero in the main specification, becomes significant in this exercise.

Table 3.D.2 – Infant mortality rate controlling for the incidence of fire

	Infant mortality rate (by thousand live births)						
	1d	1w	1m	3m	1y	Female 1y	Male 1y
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Preg. coal exp.	0.002**	0.0003	0.003***	0.003*	0.003*	0.008***	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)
Preg. n. gas exp.	0.0001	−0.0004	0.0002	−0.00004	0.0004	0.001	−0.001
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	−0.001	0.0001	0.001	−0.0002	−0.001	−0.003	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Preg. bio exp.	−0.004	0.009	0.012	0.004	−0.002	0.006	−0.015*
	(0.010)	(0.013)	(0.013)	(0.010)	(0.009)	(0.016)	(0.008)
Mean dep. var	2.65	6.3	8.39	10.66	12.12	11.1	13.05
Observations	35,701	35,701	35,701	35,701	35,701	34,970	35,043
Adjusted R^2	0.030	0.032	0.031	0.034	0.031	0.007	−0.002

Notes: *p$<$0.1; **p$<$0.05; ***p$<$0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for only female, and only male deaths in the first year of life. Main explanatory variables are the exposure to each type of power plant during the pregnancy period. In this specification we further add another variable to control for nearby fire spots.

3.E Infant mortality by trimester of exposure

Table 3.E.1 divides exposure in each trimester of pregnancy. The dependent variables are similar to the main specification presented in Table 3.5.1, with columns 1 through 5 representing infant mortality in the first day, first week, first month, first three months, and first year of life. Columns 6 and 7 are infant mortality rates up to 1 year of age from female and male children, and column 8 shows the results for the same potential mortality rate variable built to study fetal selection

in section 3.5.1. This variable equals fetal deaths plus infant mortality divided by fetal deaths plus live births, normalized by a thousand. A caveat arises when studying fetal death due to exposure in each trimester of pregnancy. Because we don't have precise information of the gestational length before fetal death for the full sample, if, for example, a fetus died in month six of pregnancy, what we are considering as exposure in the last trimester is actually the exposure in the second trimester. Besides that, most of the fetal deaths observed happened in the last trimester of pregnancy. As of 2012, data from SIM/Datasus started to report the fetal age in the moment of death. Figure 3.E.1 presents the distribution of deaths by fetal week, and the large majority of the deaths occur either in the second half of the second trimester or in the third trimester. There is no clear pattern of influence of the exposure in each trimester of pregnancy on the infant mortality rate.

Table 3.E.1 – Infant mortality by trimester of exposure

	Dependent variable:							
	1d	1w	1m	3m	1y	Female 1y	Male 1y	F.D. + 1y rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coal exp.: 3rd trim	0.00001	−0.001	0.002	0.0004	−0.003	0.017**	−0.001***	−0.003
	(0.002)	(0.003)	(0.003)	(0.003)	(0.009)	(0.008)	(0.0002)	(0.004)
2nd trim	0.005*	0.004**	0.004	0.005	0.013	0.005	−0.0001	0.016**
	(0.003)	(0.002)	(0.003)	(0.003)	(0.012)	(0.006)	(0.0003)	(0.008)
1st trim	0.00003	−0.003***	−0.0001	0.001	−0.008	0.007	−0.001**	0.006
	(0.001)	(0.001)	(0.002)	(0.004)	(0.012)	(0.012)	(0.0003)	(0.007)
N.gas exp.: 3rd trim	0.001**	0.0001	0.001	0.0004	0.001	−0.0002	0.0001	−0.0002
	(0.0005)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.0002)	(0.002)
2nd trim	−0.001	−0.001	−0.001	−0.001	−0.002	0.002	−0.0001	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.0002)	(0.002)
1st trim	0.0002	0.001	0.002	0.001	0.002	0.001	0.0004	−0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.0003)	(0.002)
Oil exp.: 3rd trim	−0.004***	−0.005*	−0.004	−0.004	−0.014	−0.003	−0.001*	−0.006
	(0.001)	(0.003)	(0.003)	(0.004)	(0.010)	(0.004)	(0.001)	(0.005)
2nd trim	0.00001	0.005*	0.006*	0.002	0.002	−0.002	0.001	−0.005
	(0.002)	(0.003)	(0.003)	(0.004)	(0.010)	(0.005)	(0.001)	(0.006)
1st trim	0.001	−0.001	−0.001	0.001	0.006	−0.009	−0.0004	0.008
	(0.002)	(0.002)	(0.003)	(0.004)	(0.009)	(0.006)	(0.001)	(0.006)
Bio exp.: 3rd trim	−0.007	0.014	0.008	−0.003	−0.086***	0.003	−0.001	−0.032**
	(0.017)	(0.028)	(0.031)	(0.028)	(0.019)	(0.033)	(0.002)	(0.013)
2nd trim	0.008	0.023	0.030*	0.026	0.040	0.009	−0.0002	−0.016
	(0.007)	(0.020)	(0.017)	(0.016)	(0.037)	(0.013)	(0.001)	(0.019)
1st trim	−0.014	−0.012**	−0.007*	−0.015***	−0.110***	0.050***	−0.002**	0.031**
	(0.011)	(0.006)	(0.003)	(0.004)	(0.033)	(0.014)	(0.001)	(0.015)
Mean dep. var	2.65	6.3	8.39	10.66	25.23	10.84	1.71	22.48
Observations	35,701	35,701	35,701	35,701	34,970	35,043	35,701	35,701
Adjusted R^2	0.030	0.032	0.031	0.034	0.027	0.008	0.495	0.058

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in Columns 1, 2, 3, 4, and 5 are the infant mortality rate in the 1st day (1d), 1st week (1w), 1st month (1m), 1st three months (3m), and 1st year of life (1y). Columns 6 and 7 show results for only female, and only male deaths in the first year of life. Column 8 shows results for the potential infant deaths considering fetal death. Main explanatory variables are the exposure to each type of power plant during each trimester of pregnancy.

Figure 3.E.1 – Count of fetal death by gestational week
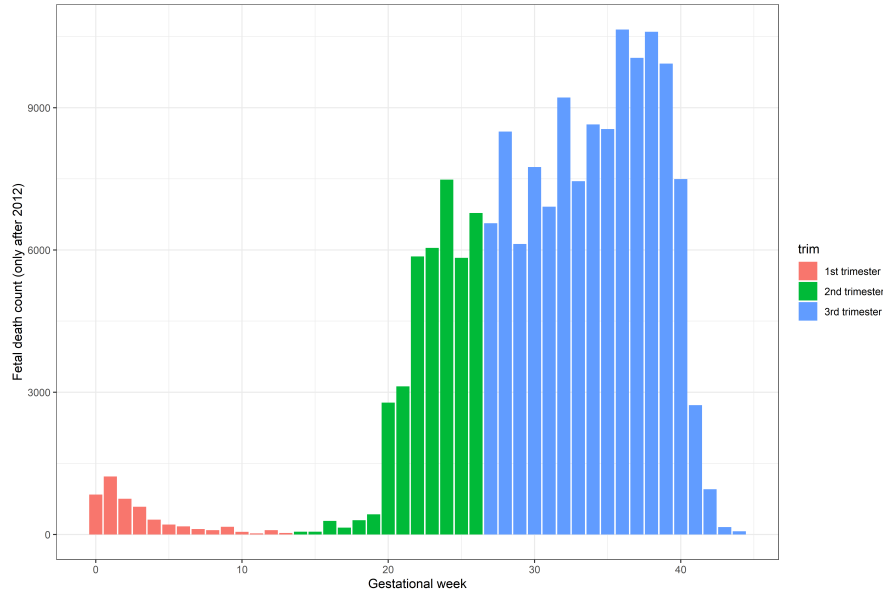
Notes: Data from Datasus/SIM. As of 2012 the fetal death data contains information on the gestational week of the fetal death. This figure counts how many deaths there are in each gestational week.

3.F Alternative definition of the exposure variable

Table 3.F.1 compares the benchmark specification (reproduced in column 3) with alternative definitions of the exposure variable. In columns 1 and 2, we consider that the whole municipality is exposed to a power plant: in column 1 if the population-weighted centroid is downwind of a power plant, and in column 2 if the center of the municipality is downwind.[3.F.1] Column 3 is the benchmark specification that considers the share of the population exposed to a power plant, based on downwind Census Tracts, as discussed in Section 3.3.3. Comparing the benchmark specification with the simpler exposure variables presented in columns 1 and 2, we notice larger coefficients with different signs in the coal variable. In column 4, instead of considering a Census Tract to calculate the share of the population exposed, we consider the population grid from the Oak Ridge National Laboratory (presented in Section 3.3.4). Like in the previous case, initially we

---

[3.F.1]Center of the municipality considered as the region with highest population density.

consider a grid to be exposed to a power plant if its centroid is downwind of a facility, and then we calculate the share of the population living in these downwind grids. Results are similar to the benchmark specification. All these initial four columns consider the main sample of municipalities, where a city is selected if its center lies within 50 km of a power plant. In column 5, we use a different sample selection and include all municipalities that have at least one Census Tract within 50 km of a power plant. If there is a large municipality far from a power plant, but with one Census Tract within the analyzed range, then this whole municipality will be included in the sample (but the exposure variable still considers only the share of the population from the Census Tracts within the 50 km range). In this last specification, results point to the same direction and are even stronger than in the benchmark definition. Once we start to consider the population exposed to the operating power plant, as of column 3, the results are fairly stable and with comparable magnitudes.

Table 3.F.1 – Infant mortality rate with different rules for municipality exposure

	Infant mortality (death by a thousand live births)				
	P. centroid	Center	CT: benchm	Grid	CT in range
	(1)	(2)	(3)	(4)	(5)
Preg. coal exp.	−0.00001*	−0.00002	0.003*	0.003*	0.005***
	(0.00001)	(0.00001)	(0.002)	(0.002)	(0.001)
Preg. n. gas exp.	0.00004*	0.00002	0.001	0.001	0.0003
	(0.00002)	(0.00002)	(0.001)	(0.001)	(0.0003)
Preg. oil exp.	−0.0001***	−0.0001**	−0.001	−0.001	−0.0005
	(0.00002)	(0.00002)	(0.001)	(0.001)	(0.002)
Preg. biomass exp.	−0.0001**	−0.0001***	−0.003	−0.002	0.004
	(0.00003)	(0.00002)	(0.010)	(0.009)	(0.039)
Mean dep. var	12.12				12.16
Observations	35,701	35,701	35,701	35,701	49,895
Adjusted R^2	0.031	0.031	0.031	0.031	0.030

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable in all columns are the infant mortality rate in the first year of life. Each column shows the result for a specific definition of the exposure variable.

3.G   Infant mortality rate heterogeneity analysis

In this section, we analyze the gender and age heterogeneities of the mortality rates by the different diseases discussed in Table 3.5.3. Each table contains the results for one specific type of disease. The columns represent the age of death, and each panel shows the results for a specific gender.

Table 3.G.1 suggests that exposure to coal leads to an increase in the mortality rate by respiratory diseases of female children but not males, while the opposite happens with exposure to natural gas power plants. These effects are only statistically significant in the specification that evaluates death up to one year of

age, and death at younger ages are all statistically equal to zero. In this analysis, we also observe an unexpected sign in the impact of biomass power plants on female deaths in the first year of life.

Table 3.G.1 – Age and gender heterogeneity for respiratory diseases

| | Mortality rate by respiratory disease | | | | |
| | 1d | 1w | 1m | 3m | 1y |
	(1)	(2)	(3)	(4)	(5)
Male children					
Preg. coal exp.	0.000	−0.00002	−0.00004	−0.001	−0.001
	(0.000)	(0.00002)	(0.0001)	(0.001)	(0.001)
Preg. n. gas exp.	0.000	0.00001	0.00003	0.0002	0.0004**
	(0.000)	(0.00001)	(0.0001)	(0.0001)	(0.0002)
Preg. oil exp.	0.000	−0.0001	−0.0002	−0.0003	−0.001
	(0.000)	(0.0001)	(0.0002)	(0.001)	(0.001)
Preg. biomass exp.	0.000	0.00004	0.0004	−0.001	0.001
	(0.000)	(0.00003)	(0.001)	(0.002)	(0.003)
Mean dep. var	0	0	0.05	0.4	0.68
Observations	35,043	35,043	35,043	35,043	35,043
Adjusted R^2		−0.033	−0.012	0.011	0.008
Female children					
Preg. coal exp.	0.00000	−0.00001	0.00004	0.001	0.001**
	(0.00001)	(0.00001)	(0.0001)	(0.001)	(0.0004)
Preg. n. gas exp.	0.00000	−0.00001	0.00005	0.0001	0.0003
	(0.00000)	(0.00001)	(0.0001)	(0.0001)	(0.0002)
Preg. oil exp.	0.00000	0.00001	−0.00004	0.0003	0.0004
	(0.00000)	(0.00001)	(0.0001)	(0.0004)	(0.001)
Preg. biomass exp.	−0.00002	−0.00003	0.0001	−0.001	−0.004***
	(0.00002)	(0.00004)	(0.0004)	(0.001)	(0.001)
Mean dep. var	0	0	0.05	0.32	0.57
Observations	34,970	34,970	34,970	34,970	34,970
Adjusted R^2	−0.042	−0.029	0.038	0.010	0.008

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable is infant death by respiratory diseases (ICD-10 codes that start with the letter "J"). Each columns analyzes a specific period of death, and each panel shows the results for a specific gender. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Tables 3.G.2 and 3.G.3 suggest that the effects observed in the mortality from conditions originated in the perinatal period are driven mainly by male children and can be observed in all ages but on the first day of life.

Table 3.G.2 – Age and gender heterogeneity for diseases originated in the perinatal period

	Mortality rate by perinatal disease				
	1d	1w	1m	3m	1y
	(1)	(2)	(3)	(4)	(5)
Male children					
Preg. coal exp.	−0.0001	−0.004	0.002	−0.0004	0.00004
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)
Preg. n. gas exp.	0.0002	−0.0003	−0.0001	−0.0004	−0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	−0.0004	0.002**	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	−0.010	−0.011	−0.002	−0.009	−0.010
	(0.008)	(0.012)	(0.010)	(0.009)	(0.009)
Mean dep. var	2.19	5.46	7.01	7.52	7.63
Observations	35,043	35,043	35,043	35,043	35,043
Adjusted R^2	0.013	0.012	0.013	0.014	0.014
Female children					
Preg. coal exp.	0.001	0.001	0.0004	0.001	0.001
	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Preg. n. gas exp.	0.0002	0.0002	0.001	0.001	0.001
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	0.0003	−0.0003	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	−0.009	0.006	0.008	0.008	0.009
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Mean dep. var	1.81	4.33	5.71	6.17	6.25
Observations	34,970	34,970	34,970	34,970	34,970
Adjusted R^2	0.013	0.016	0.016	0.016	0.017

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable is infant death related to certain conditions originated in the perinatal period (ICD-10 codes that start with the letter "P"). Each columns analyzes a specific period of death, and each panel shows the results for a specific gender. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.G.3 – Age and gender heterogeneity for respiratory and cardiovascular diseases originated in the perinatal period

| | Mortality rate by resp. card. perinatal disease | | | | |
| | 1d | 1w | 1m | 3m | 1y |
	(1)	(2)	(3)	(4)	(5)
Male children					
Preg. coal exp.	0.001	−0.0003	0.002**	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Preg. n. gas exp.	−0.00000	−0.0003	−0.0002	−0.001	−0.0004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	0.00002	0.001**	0.002***	0.002***	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	−0.002	−0.007	−0.009*	−0.010*	−0.010*
	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Mean dep. var	0.89	2.27	2.69	2.85	2.92
Observations	35,043	35,043	35,043	35,043	35,043
Adjusted R^2	0.009	0.005	0.008	0.009	0.009
Female children					
Preg. coal exp.	−0.002***	−0.002	−0.001	−0.002	−0.002
	(0.0003)	(0.001)	(0.003)	(0.003)	(0.003)
Preg. n. gas exp.	0.001	0.001	0.001*	0.001	0.001*
	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. oil exp.	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	−0.011***	−0.009***	−0.009***	−0.008***	−0.008***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Mean dep. var	0.73	1.8	2.14	2.28	2.33
Observations	34,970	34,970	34,970	34,970	34,970
Adjusted R^2	−0.003	−0.001	−0.002	−0.003	−0.003

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable is infant death related to respiratory and cardiovascular disorders specific to the perinatal period (ICD-10 codes "P20-29"). Each columns analyzes a specific period of death, and each panel shows the results for a specific gender. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.G.4 – Age and gender heterogeneity for congenital malformation respiratory and cardiovascular diseases

	Mortality rate by hart and resp malformation				
	1d	1w	1m	3m	1y
	(1)	(2)	(3)	(4)	(5)
Male children					
Preg. coal exp.	0.001	0.0002	0.0003	−0.0002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. n. gas exp.	−0.00002	−0.00004	−0.0003	−0.001	−0.001
	(0.0001)	(0.0002)	(0.0003)	(0.0004)	(0.0005)
Preg. oil exp.	0.0001	−0.00005	−0.001	−0.0004	−0.001
	(0.0002)	(0.0004)	(0.001)	(0.001)	(0.001)
Preg. biomass exp.	0.003	0.014***	0.014***	0.019***	0.019***
	(0.002)	(0.004)	(0.003)	(0.005)	(0.005)
Mean dep. var	0.17	0.52	0.81	1.17	1.4
Observations	35,043	35,043	35,043	35,043	35,043
Adjusted R^2	0.044	0.010	0.007	−0.017	−0.037
Female children					
Preg. coal exp.	−0.0004	−0.001	−0.001	−0.001	0.001
	(0.0003)	(0.001)	(0.001)	(0.001)	(0.001)
Preg. n. gas exp.	−0.0001	0.00000	0.00004	−0.0002	−0.0002
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Preg. oil exp.	−0.0002	0.0002	0.0002	−0.0004	−0.001
	(0.0002)	(0.0002)	(0.0003)	(0.001)	(0.001)
Preg. biomass exp.	−0.0004	0.002***	−0.002	−0.0003	0.003
	(0.001)	(0.001)	(0.003)	(0.001)	(0.004)
Mean dep. var	0.12	0.4	0.64	0.95	1.18
Observations	34,970	34,970	34,970	34,970	34,970
Adjusted R^2	0.010	−0.0001	0.004	0.004	−0.001

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the average number of live births in each municipality. Standard errors clustered at the municipality block level. The dependent variable is infant death related to congenital malformations (MF), deformations and chromosomal abnormalities of the respiratory and circulatory systems (ICD-10 codes "Q20-28" and "Q30-34"). Each columns analyzes a specific period of death, and each panel shows the results for a specific gender. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

3.H Alternative specification for the hospitalization by respiratory diseases

Table 3.H.1 presents the results for hospitalization by respiratory diseases for all age groups using different specifications that grow in complexity until, in column 8, it reaches the preferred model. Columns 1 and 2 are, respectively, the regressions without any controls and fixed effects, and including only the weather controls. In these specifications, exposure to coal and biomass power plants seem to have a positive effect on hospitalization. In the following two columns, we include CT and month fixed effects. As expected, the explanatory power increases, but the coefficients become negative and statistically significant. In column 5, we include a month fixed effect for each municipality, allowing non-parametric trends to differ in each city. In this specification, we are capturing differences between the exposure of Census Tracts within a specific municipality and month. In columns 6, 7, and 8, we add, respectively, fixed effects for block of municipalities by calendar-month, control variables for predominant wind direction, and both of these variables together. Once we include the municipality by month fixed effect, results are mostly indistinguishable from zero.

Table 3.H.1 – Hospitalization by all respiratory diseases with different specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coal expo	0.00264***	0.00225**	-0.00045**	-0.00037**	-0.00041	-0.00041	-0.00028	-0.00028
	(0.00098)	(0.00097)	(0.00019)	(0.00019)	(0.00028)	(0.00028)	(0.00026)	(0.00026)
Ngas expo	-0.00029	-0.00037	-0.00003	-0.00004	-0.00015	-0.00015	-0.00016	-0.00016
	(0.00029)	(0.00028)	(0.00009)	(0.00009)	(0.00022)	(0.00022)	(0.00023)	(0.00023)
Oil expo	-0.00101*	0.00000	-0.00058**	-0.00045*	-0.00065*	-0.00065*	-0.00066*	-0.00066*
	(0.00058)	(0.00053)	(0.00024)	(0.00024)	(0.00036)	(0.00036)	(0.00036)	(0.00036)
Bio expo	0.00399*	0.00504**	-0.00443***	-0.00386**	0.00163	0.00163	0.00149	0.00149
	(0.00222)	(0.00207)	(0.00168)	(0.00176)	(0.00130)	(0.00130)	(0.00129)	(0.00129)
Mean dep.var	0.37	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.00	0.00	0.76	0.76	0.77	0.77	0.77	0.77
Weather ctrl		x	x	x	x	x	x	x
CT FE			x	x	x	x	x	x
Month FE				x				
Month by mun FE					x	x	x	x
Cal. mth by block mun. FE						x		x
Wind pttrn. ctrl							x	x

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the number of inhabitants in each Census Tract. Standard-errors clustered at municipality level. The dependent variable in all columns is the hospitalization by all types of respiratory diseases considering every age group. Each column shows the results for one specification determined by the controls' panel on the bottom of the table. Main explanatory variables are the exposure to each type of power plant during a month.

3.I Age and disease heterogeneity analysis for hospitalization by respiratory disease

This section analyzes age heterogeneity for the hospitalization by different diseases. The results are similar to that observed in Table 3.5.7, with most coefficients indistinguishable from zero.

Table 3.I.1 – Hospitalization by asthma in different age groups

	All ages (1)	Age 0-1 (2)	Age 0-1: male (3)	Age 0-1: female (4)	Age 0-5 (5)	Age 6-64 (6)	Age 65+ (7)
Coal expo	-0.00002	-0.00069	-0.00053	-0.00016	-0.00014	-0.00002	0.00022**
	(0.00005)	(0.00090)	(0.00067)	(0.00044)	(0.00020)	(0.00004)	(0.00010)
Ngas expo	-0.00005	-0.00005	0.00003	-0.00008	-0.00011	-0.00004	-0.00012
	(0.00006)	(0.00056)	(0.00033)	(0.00028)	(0.00022)	(0.00006)	(0.00010)
Oil expo	-0.00011	0.00015	-0.00029	0.00044	-0.00102*	-0.00004	0.00002
	(0.00008)	(0.00077)	(0.00072)	(0.00041)	(0.00055)	(0.00004)	(0.00012)
Bio expo	0.00012	0.00440*	0.00342	0.00098	0.00151*	0.00005	-0.00011
	(0.00015)	(0.00258)	(0.00244)	(0.00148)	(0.00077)	(0.00012)	(0.00040)
Mean dep.var	0.04	0.44	0.28	0.17	0.26	0.02	0.04
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.55	0.24	0.19	0.13	0.31	0.58	0.27

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the number of inhabitants in each Census Tract in each age group. Standard-errors clustered at municipality level. Each column shows the result for asthma hospitalization in a specific age group. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.I.2 – Hospitalization by acute respiratory diseases in different age groups

	All ages (1)	Age 0-1 (2)	Age 0-1:male (3)	Age 0-1: female (4)	Age 0-5 (5)	Age 6-64 (6)	Age 65+ (7)
Coal expo	-0.00002	-0.00443	-0.00231	-0.00211	-0.00029	-0.00001	0.00014
	(0.00003)	(0.00449)	(0.00252)	(0.00227)	(0.00038)	(0.00003)	(0.00012)
Ngas expo	0.00003	-0.00059	0.00004	-0.00063	-0.00012	0.00004	0.00004
	(0.00003)	(0.00208)	(0.00134)	(0.00083)	(0.00027)	(0.00004)	(0.00004)
Oil expo	-0.00015**	-0.00266	-0.00273*	0.00008	-0.00114*	-0.00005*	-0.00005
	(0.00007)	(0.00231)	(0.00150)	(0.00101)	(0.00062)	(0.00003)	(0.00006)
Bio expo	-0.00004	0.00379	0.00409	-0.00030	0.00047	-0.00007	0.00019
	(0.00013)	(0.00387)	(0.00346)	(0.00205)	(0.00079)	(0.00007)	(0.00029)
Mean dep.var	0.04	2.04	1.22	0.82	0.38	0.01	0.02
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.46	0.42	0.36	0.29	0.45	0.34	0.15

Notes: *p<0.1; **p<0.05; ***p<0.01. Regressions are weighted by the number of inhabitants in each Census Tract in each age group. Standard-errors clustered at municipality level. Each column shows the result for hospitalization by acute respiratory diseases in a specific age group. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.I.3 – Hospitalization by COPD diseases in different age groups

	All ages (1)	Age 0-1 (2)	Age 0-1:male (3)	Age 0-1: female (4)	Age 0-5 (5)	Age 6-64 (6)	Age 65+ (7)
Coal expo	-0.00000	-0.00019	0.00010	-0.00030	0.00003	0.00001	-0.00111
	(0.00006)	(0.00056)	(0.00029)	(0.00044)	(0.00013)	(0.00005)	(0.00077)
Ngas expo	0.00000	-0.00009	-0.00012	0.00004	-0.00003	0.00000	0.00002
	(0.00002)	(0.00021)	(0.00016)	(0.00010)	(0.00005)	(0.00002)	(0.00013)
Oil expo	-0.00002	-0.00021	-0.00015	-0.00006	-0.00006	-0.00001	-0.00011
	(0.00002)	(0.00023)	(0.00017)	(0.00012)	(0.00009)	(0.00001)	(0.00020)
Bio expo	-0.00046	0.00050	-0.00115	0.00165	-0.00002	-0.00025	-0.00395
	(0.00047)	(0.00122)	(0.00089)	(0.00117)	(0.00023)	(0.00021)	(0.00400)
Mean dep.var	0.03	0.14	0.08	0.06	0.05	0.02	0.21
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.61	0.11	0.08	0.06	0.16	0.50	0.38

Notes: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Regressions are weighted by the number of inhabitants in each Census Tract in each age group. Standard-errors clustered at municipality level. Each column shows the result for hospitalization by COPD (Chronic Obstructive Pulmonary Disease) in a specific age group. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

Table 3.I.4 – Hospitalization by pneumonia diseases in different age groups

	All ages (1)	Age 0-1 (2)	Age 0-1:male (3)	Age 0-1: female (4)	Age 0-5 (5)	Age 6-64 (6)	Age 65+ (7)
Coal expo	-0.00020	0.00025	-0.00100	0.00125	-0.00076	-0.00012	-0.00233
	(0.00020)	(0.00610)	(0.00386)	(0.00297)	(0.00156)	(0.00013)	(0.00224)
Ngas expo	-0.00014	-0.00199	-0.00026	-0.00173	-0.00059	-0.00009	-0.00024
	(0.00017)	(0.00245)	(0.00161)	(0.00119)	(0.00054)	(0.00015)	(0.00035)
Oil expo	-0.00039*	-0.00930**	-0.00529**	-0.00401**	-0.00294**	-0.00012	-0.00051
	(0.00022)	(0.00431)	(0.00269)	(0.00198)	(0.00131)	(0.00010)	(0.00059)
Bio expo	0.00130	0.01186	0.00749	0.00437	0.00556	0.00046	0.00393
	(0.00118)	(0.01592)	(0.01021)	(0.00620)	(0.00843)	(0.00042)	(0.00579)
Mean dep.var	0.19	3.26	1.88	1.38	1.06	0.08	0.63
Obs	3451308	3451308	3451308	3451308	3451308	3451308	3451308
R-squared	0.70	0.46	0.40	0.36	0.56	0.60	0.49

Notes: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Regressions are weighted by the number of inhabitants in each Census Tract in each age group. Standard-errors clustered at municipality level. Each column shows the result for pneumonia hospitalization in a specific age group. Main explanatory variables are the exposure to each type of power plant during the pregnancy period.

4 Conclusion

In this dissertation, we discussed health consequences that could arise in societies relying on infrastructures not adapted to the new scenarios of a changing climate. In the first study, we analyzed a severe and unexpected drought that affected the Metropolitan Region of São Paulo in 2014 and 2015. To cope with the water scarcity, the local water company adopted a series of strict measures with potential unintended health consequences to the population. We found that, although important to save water, the policy of reducing the pressure of water within the distribution network had a negative effect on the hospitalization rates. Analyzing variations in bottled water prices, we also argued that the population engaged with defensive behavior against the water issues and that this avoidance behavior influenced the observed results. Although our results suggested that both the pressure reduction policy and the avoidance behavior affected hospitalization rates, we can not fully disentangle their effects.

In the second study from this dissertation, we investigated the effect of using thermoelectric power plants moved by different fuels on health indicators of nearby communities. The results suggested that using coal-fueled power plants increases the infant mortality rate, especially for female children. We also found evidence that coal-fueled power plants influence the incidence of fetal deaths, therefore, the observed results on infant mortality rate are likely underestimated.

These studies suggest that abrupt variations in water availability can influence health outcomes of societies. Because adapting the infrastructure to a new situation of water scarcity is costly and time-consuming, an increase in the frequency and intensity of droughts can affect the health of the population, posing a challenge to regions pursuing improvements in the Sustainable Development Goals.

REFERENCES

ALMOND, D.; CURRIE, J. Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*, v. 25, n. 3, p. 153–172, 2011.

ANTUNES, R. A Economia da Região Metropolitana de São Paulo. *Boletim Informações Fipe*, Abril, p. 28–33, 2018. Disponível em: <https://downloads.fipe.org.br/publicacoes/bif/bif451-28-33.pdf>.

ARCEO, E.; HANNA, R.; OLIVA, P. Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City. *Economic Journal*, v. 126, n. 591, p. 257–280, 2016.

ASHRAF, N. et al. Water, Health and Wealth. 2017. Disponível em: <http://www.nber.org/papers/w23807>.

AUFFHAMMER, M. et al. Using Weather Data and Climate Model Output in Economic Analyses of Climate Change. *Review of Environmental Economics and Policy*, v. 7, n. 2, p. 181–198, 2013.

AVELINO, A. F.; HEWINGS, G. J.; GUILHOTO, J. *EPSIM - A Social-Environmental Regional Sequential Interindustry Economic Model for Energy Planning: Evaluating the Impacts of New Power Plants in Brazil*. 2014. 1–29 p. Disponível em: <https://ssrn.com/abstract=2402277>.

BAI, J.; PERRON, P. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, v. 18, n. 1, p. 1–22, 2003. ISSN 08837252.

BURKHARDT, J. et al. The effect of pollution on crime: Evidence from data on particulate matter and ozone. *Journal of Environmental Economics and Management*, Elsevier Inc., v. 98, p. 1–19, 2019.

CHAGAS, A. L.; AZZONI, C. R.; ALMEIDA, A. N. A spatial difference-in-differences analysis of the impact of sugarcane production on respiratory diseases. *Regional Science and Urban Economics*, Elsevier B.V., v. 59, p. 24–36, 2016.

CHAY, K. Y.; GREENSTONE, M. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, v. 118, n. 3, p. 1121–1167, 2003.

CHEN, S.; GUO, C.; HUANG, X. Air Pollution, Student Health, and School Absences: Evidence from China. *Journal of Environmental Economics and Management*, Elsevier Ltd, v. 92, p. 465–497, 2018.

134

COELHO, C. A.; CARDOSO, D. H.; FIRPO, M. A. Precipitation diagnostics of an exceptionally dry event in São Paulo, Brazil. *Theoretical and Applied Climatology*, Theoretical and Applied Climatology, v. 125, n. 3-4, p. 769–784, 2016.

CONAMA. *Resolution no 003/1990*. Brasilia, DF: CONAMA, 1990. Disponível em: <http://www2.mma.gov.br/port/conama/res/res90/res0390.html>.

CURRIE, J. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, v. 47, n. 1, p. 87–122, 2009.

CURRIE, J. et al. Something in the Water: Contaminated Drinking Water and Infant Health. *Canadian Journal of Economics*, v. 46, n. 3, p. 791–810, 2013.

CURRIE, J.; NEIDELL, M. Air pollution and infant health: What Can We Learn From California's Recent Experience? *The Quarterly Journal of Economics*, v. 120, n. 3, p. 1003–1030, 2005.

CURRIE, J.; NEIDELL, M.; SCHMIEDER, J. F. Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics*, v. 28, n. 3, p. 688–703, 2009.

CURRIE, J.; WALKER, R. Traffic congestion and infant health: Evidence from E-Zpass. *American Economic Journal: Applied Economics*, v. 3, n. 1, p. 65–90, 2011.

CURRIE, J. et al. What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution? *Annual Review of Resource Economics*, v. 6, p. 217–247, 2014.

DERYUGINA, T. et al. The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review*, v. 109, n. 12, p. 4178–4219, 2019.

DESCHÊNES, O.; GREENSTONE, M.; GURYAN, J. Climate Change and Birth Weight. *American Economic Review: Papers & Proceedings*, v. 99, n. 2, p. 211–217, 2009.

EPE. *Plano Nacional de Energia 2030: Geração termelétrica - Biomassa*. Rio de Janeiro, RJ, 2007. 250 p.

EPE. *Plano Nacional de Energia 2030: Geração Termelétrica - Carvão Mineral*. Rio de Janeiro, RJ, 2007. 146 p.

EPE. *Balanço Energético Nacional 2010: Ano Base 2009.* Rio de Janeiro, RJ, 2010. 276 p.

EPE. *Balanço Energético Nacional 2018: Ano Base 2017.* Rio de Janeiro, RJ, 2018. 294 p.

EPE. *Plano Decenal de Expansão de Energia 2027.* Rio de Janeiro, RJ, 2018. 345 p.

EYER, J.; WICHMAN, C. J. Does water scarcity shift the electricity generation mix toward fossil fuels? Empirical evidence from the United States. *Journal of Environmental Economics and Management,* Elsevier Inc., v. 87, p. 224–241, 2018.

FONTANAZZA, C. M. et al. Contaminant intrusion through leaks in water distribution system: experimental analysis. *Procedia Engineering,* Elsevier B.V., v. 119, p. 426–433, 2015.

FUNASA. *Impactos Na Saúde e no Sistema Único de Saúde Decorrentes de Agravos Relacionados a um Saneamenteo Ambiental Inadequado.* Brasília, DF, 2010. 248 p.

GALIANI, S.; GERTLER, P.; SCHARGRODSKY, E. Water for Life: The Impact of the Privatization of Water Services on Child Mortality. *Journal of Political Economy,* v. 113, n. 1, p. 83–120, 2005.

GAMPER-RABINDRAN, S.; KHAN, S.; TIMMINS, C. The impact of piped water provision on infant mortality in Brazil : A quantile panel data approach. *Journal of Development Economics,* Elsevier B.V., v. 92, n. 2, p. 188–200, 2010.

GEHRSITZ, M. The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management,* Elsevier Inc., v. 83, p. 121–144, 2017.

Graff Zivin, J.; NEIDELL, M. The Effect of Pollution on Worker Productivity. *American Economic Review,* v. 102, n. 7, p. 3652–3673, 2012.

Graff Zivin, J.; NEIDELL, M.; SCHLENKER, W. Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption. *American Economic Review: Papers & Proceedings,* v. 101, n. 3, p. 448–453, 2011.

GREENSTONE, M.; HANNA, R. Environmental Regulations, Air and Water Pollution, and Infant Mortality in India. *American Economic Review,* v. 104, n. 10, p. 3038–3072, 2014.

GROVER, D.; LUCINDA, C. R. An Evaluation of the Policy Response to Drought in the City of São Paulo , Brazil: An Election Cycle Interpretation of Effectiveness. *The Journal of Development Studies*, p. 1–18, 2020.

GUIDETTI, B.; PEREDA, P.; SEVERNINI, E. Health Shocks under Hospital Capacity Constraint: Evidence from Air Pollution in Sao Paulo, Brazil. 2020. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3596677>.

HANNA, R.; OLIVA, P. The Effect of Pollution on Labor Supply: Evidence from Mexico City. *Journal of Public Economics*, v. 122, p. 68–79, 2015.

HE, G.; FAN, M.; ZHOU, M. The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games. *Journal of Environmental Economics and Management*, Elsevier, v. 79, p. 18–39, 2016.

HINRICHS, R. A.; KLEINBACH, M. Air Pollution and Energy Use. In: *Energy: Its Use and the Environment*. Fifth edit. Boston, MA: Brooks/Cole CENGAGE Learning, 2012. cap. 8.

IBGE. Notas técnicas do Censo de 2010. 2010.

IPCC. *Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press, 2013. 1535 p.

IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York, NY: Cambridge University Press, 2014. 1150 p.

ISEN, A.; ROSSIN-SLATER, M.; WALKER, W. R. Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy*, v. 125, n. 3, p. 848–902, 2017.

JALAN, J.; RAVALLION, M. Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics*, v. 112, n. 1, p. 153–173, 2003.

JAYACHANDRAN, S. Air quality and early-life mortality: Evidence from Indonesia's wildfires. *The Journal of Human Resources*, v. 44, n. 4, p. 916–954, 2009.

JEDRYCHOWSKI, W. et al. Gender differences in fetal growth of newborns exposed prenatally to airborne fine particulate matter. *Environmental Research*, Elsevier Inc., v. 109, n. 4, p. 447–456, 2009.

KNITTEL, C. R.; MILLER, D. L.; SANDERS, N. J. Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. *The Review of Economics and Statistics*, v. 98, n. 2, p. 350–366, 2016.

KOPPENSTEINER, M. F.; MANACORDA, M. Violence and birth outcomes: Evidence from homicides in Brazil. *Journal of Development Economics*, Elsevier B.V., v. 119, p. 16–33, 2016.

LÓPEZ, M. T. et al. Health impacts from power plant emissions in Mexico. *Atmospheric Environment*, v. 39, p. 1199–1209, 2005.

LUECHINGER, S. Air pollution and infant mortality: A natural experiment from power plant desulfurization. *Journal of Health Economics*, Elsevier B.V., v. 37, n. 1, p. 219–231, 2014.

MANGYO, E. The effect of water accessibility on child health in China. *Journal of Health Economics*, v. 27, p. 1343–1356, 2008.

MASSETTI, E. et al. *Environmental Quality and the U.S. Power Sector: Air Quality, Water Quality, Land Use and Environmental Justice*. [S.l.], 2017. 169 p.

MENDES, L. F. R.; STHEL, M. S. Thermoelectric Power Plant for Compensation of Hydrological Cycle Change: Environmental Impacts in Brazil. *Case Studies in the Environment*, p. 1–7, 2017.

MORETTI, E.; NEIDELL, M. Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. *Journal of Human Resources*, v. 46, n. 1, p. 154–175, 2011.

MULLER, N. Z.; MENDELSOHN, R. Measuring the damages of air pollution in the United States. *Journal of Environmental Economics and Management*, v. 54, n. 1, p. 1–14, 2007.

NEIDELL, M. Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations. *Journal of Human Resources*, v. 44, n. 2, p. 450–478, 2009.

ORDONEZ, P. Power Plants, Air Pollution, and Health in Colombia. 2020.

PERSSON, P.; ROSSIN-SLATER, M. Family Ruptures, Stress, and the Mental Health of the Next Generation. *American Economic Review*, v. 108, n. 4-5, p. 1253–1255, 2018.

PONTES, R. J. et al. Vector densities that potentiate dengue outbreaks in a Brazilian city. *The American Journal of Tropical Medicine and Hygiene*, v. 62, n. 3, p. 378–383, 2000.

RANGEL, M. A.; VOGL, T. S. Agricultural Fires and Health at Birth. *Review of Economics and Statistics*, v. 101, n. 4, p. 616–630, 2019.

ROCHA, R.; SOARES, R. R. Water scarcity and birth outcomes in the Brazilian semiarid. *Journal of Development Economics*, v. 112, p. 72–91, 2015.

SABESP. *Estudo de Concepção e Projeto Básico do Sistema Produtor São Lourenço. Relatório de Impacto ao Meio Ambiente - RIMA*. São Paulo, SP, 2011. 1–189 p.

SABESP. *Crise Hídrica, Estratégia e Soluções da SABESP*. São Paulo, SP, 2015. 1–95 p.

SABESP. *Relatório de Sustentabilidade*. São Paulo, SP, 2015. 1–109 p.

SANDERS, N. J.; STOECKER, C. Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics*, Elsevier B.V., v. 41, p. 30–45, 2015.

SCHLENKER, W.; WALKER, W. R. Airports, air pollution, and contemporaneous health. *Review of Economic Studies*, v. 83, n. 2, p. 768–809, 2016.

SEVERNINI, E. Impacts of nuclear plant shutdown on coal-fired power generation and infant health in the Tennessee Valley in the 1980s. *Nature Energy*, v. 2, p. 1–9, 2017.

SICILIANO, B. et al. The updated Brazilian national air quality standards: A critical review. *Journal of the Brazilian Chemical Society*, v. 31, n. 3, p. 523–535, 2020.

SONG, J. et al. Acute effects of ambient air pollution on outpatient children with respiratory diseases in Shijiazhuang, China. *BMC Pulmonary Medicine*, BMC Pulmonary Medicine, v. 18, n. 150, p. 1–10, 2018.

STANKE, C. et al. Health Effects of Drought: A Systematic Review of the Evidence. *PLoS Currents*, p. 1–35, 2013.

TANAKA, S. Environmental regulations on air pollution in China and their impact on infant mortality. *Journal of Health Economics*, Elsevier B.V., v. 42, p. 90–103, 2015.

TAO, Y. et al. Air pollution and hospital admissions for respiratory diseases in Lanzhou, China. *Environmental Pollution*, Elsevier Ltd, v. 185, p. 196–201, 2014.

TEIXEIRA, J. A. M. et al. Mortalidade no primeiro dia de vida: tendências, causas de óbito e evitabilidade em oito Unidades da Federação brasileira, entre 2010 e 2015. *Epidemiologia e Serviços de Saúde*, v. 28, n. 1, p. 1–11, 2019.

TOLMASQUIM, M. T. *Energia Termelétrica: gás natural, biomassa, carvão, nuclear*. Rio de Janeiro, RJ: Empresa de Pesquisa Energética, 2016. 417 p.

UNICEF. *Levels & Trends in Child Mortality: Report 2019*. [S.l.], 2019. 1–32 p.

United Nations. *The Sustainable Development Goals Report*. New York, NY, 2018. 40 p. Disponível em: <https://unstats.un.org/sdgs/report/2018>.

VRIJHEID, M. et al. Ambient air pollution and risk of congenital anomalies: A systematic review and meta-analysis. *Environmental Health Perspectives*, v. 119, n. 5, p. 598–606, 2011.

WHO. *WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide*. 2006. 1–22 p.

YANG, M.; CHOU, S. Y. The impact of environmental regulation on fetal health: Evidence from the shutdown of a coal-fired power plant located upwind of New Jersey. *Journal of Environmental Economics and Management*, Elsevier Ltd, v. 90, p. 269–293, 2018.

ZEILEIS, A. et al. Testing and dating of structural changes in practice. *Computational statistics and Data Analysis*, v. 44, p. 109–123, 2003.

ZHANG, J. The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China. *Journal of Health Economics*, v. 31, p. 122–134, 2012.

ZHANG, J.; XU, L. C. The long-run effects of treated water on education: The rural drinking water program in China. *Journal of Development Economics*, v. 122, p. 1–15, 2016.