Grantham Research Institute on Climate Change and the Environment

Centre for Climate Change Economics and Policy

Carbon pricing with regressive co-benefits: evidence from British Columbia's carbon tax

Lorenzo Sileci

December 2023

Centre for Climate Change Economics and Policy Working Paper No. 430
ISSN 2515-5709 (Online)

Grantham Research Institute on Climate Change and the Environment Working Paper No. 405
ISSN 2515-5717 (Online)

Economic and Social Research Council

Grantham Foundation for the Protection of the Environment

UNIVERSITY OF LEEDS

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council. Its third phase started in October 2018 with seven projects:

1. Low-carbon, climate-resilient cities
2. Sustainable infrastructure finance
3. Low-carbon industrial strategies in challenging contexts
4. Integrating climate and development policies for 'climate compatible development'
5. Competitiveness in the low-carbon economy
6. Incentives for behaviour change
7. Climate information for adaptation

More information about CCCEP is available at www.cccep.ac.uk

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and a number of other sources. It has 13 broad research areas:

1. Biodiversity
2. Climate change adaptation and resilience
3. Climate change governance, legislation and litigation
4. Climate, health and environment
5. Environmental behaviour
6. Environmental economic theory
7. Environmental policy evaluation
8. International climate politics
9. Science and impacts of climate change
10. Sustainable public and private finance
11. Sustainable natural resources
12. Transition to zero emissions growth
13. UK national and local climate policies

More information about the Grantham Research Institute is available at: www.lse.ac.uk/GranthamInstitute

Suggested citation:
Sileci L (2023) *Carbon pricing with regressive co-benefits: evidence from British Columbia's carbon tax*. Centre for Climate Change Economics and Policy Working Paper 430/Grantham Research Institute on Climate Change and the Environment Working Paper 405. London: London School of Economics and Political Science

Carbon Pricing with Regressive Co-benefits:

Evidence from British Columbia's Carbon Tax

Lorenzo Sileci[a,b]

[a] *Grantham Research Institute for Climate Change and the Environment*
[b] *LSE, Department of Geography and Environment*

November 2023

Abstract

I assess the air quality and environmental equity impacts of the 2008 carbon tax in British Columbia. Using high-resolution data and a synthetic difference-in-differences strategy, I find that the carbon tax has reduced $PM_{2.5}$ emissions by 5.2-10.9%. This result is heterogeneously distributed, with larger reductions in areas with lower baseline pollution, lower population density, lower material deprivation, and higher income. While all areas experience substantial positive co-benefits in terms of reduced air pollution hazard rates, quantified at \$198 per capita, my results imply a widening of the pre-existing environmental justice gaps. This dynamic represents an additional dimension of carbon tax regressiveness.

Keywords:
Carbon Tax, Air Quality, $PM_{2.5}$, Co-benefits, Environmental Justice
JEL: Q58, Q53, H23

---

* *Email:* `l.sileci@lse.ac.uk`

Acknowledgements: For helpful discussions and comments I thank, in alphabetical order: Eugenie Dugoua, Paul Ferraro, Ben Groom, Jeffrey Pagel, Charles Palmer, Elena Perra, Sefi Roth, Gregor Singer, and Hendrik Wolff, as well as participants to AERE 2023 Annual Conference, EAERE 2023 Summer Conference, envecon 2023, BIOECON 2023, and research seminars hosted by the Department of Geography and Environment at the LSE in 2022 and 2023. The author acknowledges funding from the UK Economic and Social Research Council (ESRC). The usual disclaimer applies.

1. Introduction

The major sources of CO_2 emissions are the fossil fuel combustion processes which also give rise to emissions of air pollutants. Climate change and air pollution can then be categorised as complementary global and local externalities from fossil fuel use. Therefore, efforts to control CO_2 emissions by internalising the social cost of carbon are bound to give rise to significant "co-benefits" in terms of air quality improvements.

Given the relative scarcity of long-tenured carbon pricing schemes, it is unsurprising that empirical evidence of their causal impact on local air pollution co-benefits is sporadic, and mostly limited to cap-and-trade schemes (Wagner and Preux, 2016; Liu *et al.*, 2021; Zhu *et al.*, 2022; Hernandez-Cortes and Meng, 2023) with fewer studies focussing on fuel taxes (e.g. Basaglia *et al.*, 2023). On the contrary, there is a large and growing literature which, using theoretical insights (Parry *et al.*, 2015) and simulation models (Knittel and Sandler, 2011; Zhang *et al.*, 2021), has attempted to calculate the monetary value of air pollution improvements due to carbon taxation and compare them with the cost of mitigation policies. In particular, net health co-benefits arising from carbon taxation are theorised to reach a high enough magnitude to partially or fully offset the mitigation costs for households at a national (Li *et al.*, 2018; Shindell *et al.*, 2016) and global (West *et al.*, 2013; Vandyck *et al.*, 2018) level, and may provide strong additional incentives for a swift transition to a low-carbon economy. Moreover, reductions in morbidity and mortality due to improvements in air quality are likely not to capture the full extent of the local pollution externality: a large body of research has linked air pollution to non-health outcomes (see Aguilar-Gomez *et al.*, 2022, for a review)[1], suggesting that any attempt at quantifying the monetary impact of co-benefits based on health outcomes alone would, at best, provide a lower bound of the beneficial consequences of air quality improvements.

In light of the considerable size of air pollution co-benefits, it is fundamental to examine how carbon pricing policies impact the spatial distribution of pollutants over affected populations. Higher reductions in co-pollutants induced by carbon taxation are indeed expected to arise from polluters with lower marginal abatement costs, but this efficiency criterion is blind to equity considerations (Hernandez-Cortes and

[1]Studies have linked air pollution to negative educational outcomes (Ebenstein *et al.*, 2016; Wen and Burke, 2022), increase in crime rates (Bondy *et al.*, 2020) and suicides (Persico and Marcotte, 2022), reductions in labour productivity (Graff Zivin and Neidell, 2012), and in housing prices (Sager and Singer, 2022; Freeman *et al.*, 2019).

Meng, 2023). A substantial body of research has indeed documented geographic and socio-economic disparities[2] in pollution levels within cities (e.g. Jbaily *et al.*, 2022; Currie *et al.*, 2020). It is thus paramount to inspect whether carbon taxation presents efficiency-equity trade-offs in the distribution of realised co-benefits, evidence of which is mixed in the environmental economics literature (Sheriff, 2023; Shapiro and Walker, 2021; Boyce and Pastor, 2013; Fowlie *et al.*, 2012; Grainger and Ruangmas, 2018).

In this paper, I assess the air quality co-benefits and environmental justice implications of the 2008 British Columbia carbon tax. The tax, covering approximately 75% of the Canadian province's CO2 emissions, was initially introduced at a rate of $10/tCO2, and sequentially ramped up by $5 per year until 2012, when it was frozen at $30/tCO2 until 2018. Importantly, no other Canadian Province introduced carbon pricing schemes between 2008 and 2018, when the tax was rolled out on a federal basis, which allows me to rely on a control pool drawn from other Canadian provinces. I leverage high-resolution data on $PM_{2.5}$, based on a combination of satellite observations, geo-chemical models and ground-based monitoring stations, from Meng *et al.* (2019) and van Donkelaar *et al.* (2019), and combine them with highly disaggregated socio-economic data at the Dissemination Area level[3], retrieved from the Canadian Census at 5-year intervals between 2001 and 2016. I exploit this granular dataset to assess the effect of the carbon tax on air pollution co-benefits and the dynamics of the environmental justice gap.

The central result of the paper is that the 2008 British Columbian carbon tax has resulted in statistically significant reductions in $PM_{2.5}$ concentrations, with a lower bound average estimate of -0.36 $\mu g/m^3$ and an upper bound average estimate of -0.89 $\mu g/m^3$, corresponding to a 5.2-10.9% reduction in particulate matter concentrations with respect to pre-treatment average levels. Importantly, as in e.g. Andersson (2019), Sager and Singer (2022), and Basaglia *et al.* (2023), this result is obtained by moving away from traditional two-way fixed effects difference-in-differences (TWFE-DID) estimation, in light of a violation of the foundational parallel trends assumption: particulate matter trends between British Columbian and control Dissemination Areas diverge prior to the implementation of the carbon tax, thereby biasing DID estimates. I rely on a family of estimators related to the synthetic control method (SCM) for comparative case studies (Abadie and Gardeazabal, 2003; Abadie, 2021),

[2]Also referred to as the "environmental justice gap" (Hernandez-Cortes and Meng, 2023).
[3]Corresponding roughly to US Census tracts.

employing in particular the synthetic difference-in-differences (SDID) estimator by Arkhangelsky *et al.* (2021) as my preferred methodology.

I subsequently inspect the efficiency-equity trade off, examining whether air pollution reductions arise heterogeneously within British Columbian metropolitan areas. I split the pool of treated units in quintiles of pre-existing pollution, population density, median income levels and material deprivation index and estimate the impact of the tax on $PM_{2.5}$ reductions for each quintile of these baseline characteristics. The carbon tax appears to be regressive in the spatial dimension: reductions are 1.6-2.2 times higher in the bottom quintile of pre-treatment air pollution, population density and material deprivation compared to the top quintile, and 1.7 times higher in the top median income quintile compared to the bottom quintile.

Finally, I convert my estimates of particle pollution reductions into mortality reductions[4] and associated monetary gains, relying on the concept of the Value of a Statistical Life[5]. The median monetary health gains appear to be large, in the order of \$88-402/year per capita. The central estimate of \$198 is almost double the \$115.50 per capita Low-Income Climate Action Tax Credit, the carbon tax governmental rebate accruing to low-income individuals to mitigate the cost of carbon pricing. The total annual health gains are comparable to annual carbon tax revenues at its inception (Ministry of Finance, 2009) and amount to 40-81% of annual tax revenues at maturity (Ministry of Finance, 2013). Health gains stemming from $PM_{2.5}$ are also heterogeneous over space, with greater benefits manifesting in peri-urban areas rather than in city centres, and exhibit a positive correlation with income within metropolitan areas, corroborating the evidence on the increase in the environmental justice gap.

This paper contributes to the literature on three main fronts. First, I extend the recent evidence on the impact of carbon pricing on air pollution co-benefits, with an explicit focus on carbon taxation instead of the frequently examined cap-and-trade schemes (Hernandez-Cortes and Meng, 2023) and fuel tax increases (Basaglia *et al.*, 2023). Differently from Saberian (2017), I find positive co-benefits from the 2008 carbon tax, and expand the geographic and temporal scope of the analysis. I overcome known endogeneity problems connected with the use of sparse air quality monitors (Carozzi and Roth, 2022) by relying on two sets of remotely sensed $PM_{2.5}$ data (Meng *et al.*, 2019; van Donkelaar *et al.*, 2019) which provide full coverage of

[4]Exploiting hazard rates adapted from the environmental health and epidemiology literature (Lepeule *et al.*, 2012; Krewski *et al.*, 2009).

[5]Following Fowlie *et al.* (2019) and Carozzi and Roth (2022).

the spatial and temporal extent of my dataset. Further, I dispel the notion that the carbon tax has resulted in gasoline to diesel fuel substitution, instead highlighting expected reductions in both fuels' total demand after the tax. Moreover, by exploiting highly disaggregated census information on commute mode, I provide evidence on additional mechanisms underlying the air quality improvements: BC residents substitute high emissions trips with public transport and active commute modes following the implementation of the tax. My results are thus also consistent with the findings of Pretis (2022), who found that the 2008 carbon tax reduced CO_2 emissions in the transportation sector alone.

The second contribution regards the growing literature on the relationship between environmental policies and equity. I present the first ex post analysis of the effects of a carbon tax on the "environmental justice (EJ) gap" (Hernandez-Cortes and Meng, 2023). I find that pricing carbon, while giving rise to widespread air quality co-benefits, may do so disproportionately with respect to pre-existing levels of air pollution, income, density (Carozzi and Roth, 2022) and material deprivation. My estimates thus add a data point to the nascent literature on ex post empirical evaluation of EJ effects from climate policy, which has so far reported mixed evidence (Fowlie *et al.*, 2012; Boyce and Pastor, 2013; Grainger and Ruangmas, 2018; Shapiro and Walker, 2021; Sheriff, 2023; Hernandez-Cortes and Meng, 2023). These results call for spatially heterogeneous climate interventions (Nehiba, 2022) or for additional layered instruments aimed at internalising the congestion externality in urban centres and reducing local pollution (e.g. Pestel and Wozny, 2021; Sarmiento *et al.*, 2022; Gehrsitz, 2017), with a specific focus on policy impacts on disadvantaged communities. Other examples of incremental policies to aid carbon pricing in providing co-benefits are incentives for alternative transport modes, as low-income and disadvantaged households are relatively more cash and credit constrained.

Lastly, I contribute to the environmental policy evaluation literature in a similar vein as Sager and Singer (2022). I indeed show how the traditional TWFE-DID estimator is susceptible of producing biased estimates, due to substantially diverging pre-treatment trends across treatment and control units. I solve this concern by exploiting SCM and the newly introduced SDID estimator (Arkhangelsky *et al.*, 2021) and exploiting, unlike recent studies in environmental policy evaluation (e.g. Andersson, 2019; Leroutier, 2022; Basaglia *et al.*, 2023) a subnational level treatment and a highly granular framework. In my setting, with multiple treated units and a large number of control units to draw synthetic counterfactuals from, both the SCM and SDID perform well in addressing concerns about diverging pre-treatment trends and identify unbiased and robust estimates of the impact of the carbon tax on $PM_{2.5}$ levels, improving substantially upon traditional estimators and aggregate policy settings.

The remainder of the paper begins with a description of the carbon tax and the data sources in Section 2. In Section 3, I present the identification strategy, followed by the main results in Section 4 and mechanisms underlying them in Section 5. Section 6 shows the consistency of the main analyses to alternative specifications. Lastly, I examine environmental justice concerns in Section 7, and estimate mortality reductions and associated monetary health gains in Section 8. Section 9 concludes, and additional information is reported in the Appendix.

2. Policy Context, Data and Descriptive Statistics

2.1. The 2008 British Columbian Carbon Tax

The introduction of the British Columbia (BC) carbon tax was formally announced in the provincial budget plan in February 2008, catching the public off guard due to the unexpected nature of this move by the Liberal government (Harrison, 2012; Ahmadi *et al.*, 2022). The policy aimed to reduce emissions by a minimum of 33% below 2007 levels by 2020 (Azevedo *et al.*, 2023). Implemented on July 1, 2008, the initial tax rate was set at \$10/tonne CO_2eq and increased by \$5/tonne CO_2eq annually until it reached \$30 in 2012, establishing one of the highest carbon prices globally at the time (Murray and Rivers, 2015; Azevedo *et al.*, 2023). The carbon tax rate remained at \$30 until 2018, when it increased to \$35, with a subsequent annual increment of \$5 anticipated until it reaches \$50/tonne in 2022. The tax, applicable to all fossil fuel purchases in BC, accounts for approximately 77% of the province's total greenhouse gas (GHG) emissions, underscoring the comprehensive scope of the policy (Murray and Rivers, 2015; Rivers and Schaufele, 2015; Ahmadi *et al.*, 2022; Azevedo *et al.*, 2023). Notably, the most affected sector is transportation, which contributed to 43.9% of the province's total CO_2 levels in 2007; exemptions cover exported fuels, non-combustion GHGs (e.g. landfill methane), and emissions generated outside BC[6].

A key aspect of implementing the BC carbon tax is its commitment to revenue neutrality, serving as a crucial mechanism to secure public support and mitigate

[6]This excludes a significant portion of air transportation and non-metallic mineral manufacturing emissions. Additionally, non-fossil fuel sources like fugitive emissions and chemical processes are exempted, broadening the range of exclusions (Azevedo *et al.*, 2023).

resistance to additional taxation, a notable challenge in the execution of carbon pricing schemes (Carattini *et al.*, 2017; Carattini *et al.*, 2019)[7].

The revenue-neutral design of the tax involved returning funds to consumers and businesses through various means, including direct transfers to low-income individuals, income tax reductions, and corporate tax cuts (Murray and Rivers, 2015; Ahmadi *et al.*, 2022). In particular, the achievement of revenue neutrality in BC involves two primary mechanisms. Firstly, by initiating a 5% reduction in the bottom two income tax brackets, BC secured the lowest income tax rate in Canada for individuals earning up to $122,000. This reduction was complemented by additional measures such as the "low income climate action" tax credit and the Northern and Rural Homeowner benefit (Azevedo *et al.*, 2023)[8]. Secondly, a series of reductions were applied to the general corporate tax rate, starting at 12% in 2008 and gradually decreasing to 11%, 10.5%, and 10% in 2010 and 2011, before returning to 11% in 2014. Simultaneously, the small business corporate income tax rate decreased from 4.5% to 2.5% in 2008 (Azevedo *et al.*, 2023)[9]. According to the Budget and Fiscal Plan, the carbon tax generated approximately $1.2 billion in annual revenue since 2012 when the rate stabilized at $30/tonne CO_2eq, with around $1.4 billion returned to consumers (Ahmadi *et al.*, 2022; Azevedo *et al.*, 2023).

Given the popularity of the carbon tax, it is unsurprising that economists have conducted several analyses of its effectiveness across a range of measures. Focussing on the transport fuel market, Rivers and Schaufele (2015) and Lawley and Thivierge (2018) find 5-8% reductions in gasoline demand due to the tax implementation. Azevedo *et al.* (2023) investigate the employment response to the tax: the absence of aggregate effects masks heterogeneous impacts, with large emission-intensive firms

[7] Subsequent to the initial "Axe the tax" campaigns leading up to the 2009 provincial elections, polling data indicated a sustained increase in public approval of the tax until 2015 (Murray and Rivers, 2015). However, after 2012, there was a shift towards earmarking some revenues for specific sectors, creating a mixed system of redistribution (Murray and Rivers, 2015). Public opinion on the carbon tax was initially volatile, with campaigns against it leading up to the 2009 provincial elections, but sustained approval was observed until 2015 (Murray and Rivers, 2015). Recent studies, though, suggest that attitudes towards carbon pricing may be more influenced by partisan identities than updated information about potential rebates (Mildenberger *et al.*, 2022)

[8] The low income climate action tax credit was initially set as $100 per adult plus $30 per child, and subsequently raised to $115.50 per adult and $34.50 per child (Ministry of Finance, 2009; Ministry of Finance, 2013). The Northern and Rural Homeowner Benefit amounts to $200 but only applies to howeowners in areas outside the Capital (Victoria CMA), Greater Vancouver (Vancouver CMA) and Fraser Valley (Abbotsford CMA) regional districts. The appropriate rebate to compare to health gains is thus the low income climate action tax credit.

[9] Since 2008, various tax credits, ranging from the BC Seniors Home Renovation Tax Credit to the Film Incentive BC tax credit, have been implemented, contributing to the revenue redistribution.

negatively affected and small businesses benefitting from the policy. In terms of global pollutants, Ahmadi *et al.* (2022) detect emissions reductions in the manufacturing sector, while the multisectoral analysis of Pretis (2022) identifies significant reductions in transportation emissions with negligible effects on the remaining sectors of the economy.

2.2. Data and Descriptive Statistics

In order to analyse the effect of British Columbia's 2008 carbon tax on air quality, I assemble and process information on local pollutants' concentrations, geographic characteristics, and socio-economic dynamics from multiple sources. The observational units which I employ in the analysis are Dissemination Areas (DAs), the smallest standard geographic areas for which Canadian census data are disseminated. Since the paper is concerned with analysing the effect of carbon pricing on air quality in cities, I restrict the geographic scope of the dataset to 26 Canadian Census Metropolitan Areas (CMAs), thereby excluding rural areas and smaller towns[10]. Canadian census data is obtained from von Bergmann *et al.* (2022), while DA census boundaries are converted to common geographies based on von Bergmann (2021), and using DA administrative boundaries from the 2016 Canadian census as the target geography. My final dataset is thus comprised of 25,479 DAs observed over 19 years, from 2000 to 2018, across 26 CMAs.

The dependent variable employed in the main part of the paper is yearly average $PM_{2.5}$ concentration from Meng *et al.* (2019), which combine information from satellite-retrieved Aerosol Optical Depth with simulations and ground-based observations obtained from monitoring stations readings. I extract the mean value of yearly $PM_{2.5}$, weighted by grid-cell level population counts obtained from Rose *et al.* (2020), onto the 25,479 DAs which constitute my dataset for every year between 2000 and 2018[11]. Hence, for each DA, the dependent variable takes the form:

[10]The CMAs in the dataset are: St. John's, Halifax, Saint John, Quebec, Trois Rivieres, Sherbrooke, Montreal, Ottawa, Saguenay, Kingston, Toronto, Hamilton, St. Catharine's, Kitchener, London, Windsor, Sudbury, Thunder Bay, Winnipeg, Regina, Saskatoon, Calgary, Edmonton, Abbotsford, Vancouver, and Victoria. While the number of Canadian CMAs is 35 in the latest available census wave (2016), we only keep in the dataset those CMAs which were designated as such in the 2001 Census, in order to ensure compatibility across all waves.

[11]The resolution of the $PM_{2.5}$ raster data is 0.01°x 0.01°, while population data is available for grid cells of dimension 0.0083° x 0.0083°, implying that the population raster had to be resampled at the resolution of the $PM_{2.5}$ raster in order to be viable for use in the weighted mean calculation.

$$PM_{2.5it} = \frac{1}{N_j} \sum_{j=1}^{N} Pop_{jt} * PM_{2.5jt} \tag{1}$$

Where $j = 1, ..., N$ is the number of raster grid cells in a DA i, Pop_{jt} is the population count in grid cell j at time t, and $PM_{2.5jt}$ is the value of the particulate matter raster in grid cell j at time t.

The main advantage of this source compared to data obtained from monitoring stations only (Saberian, 2017), is their much wider spatial and temporal coverage, which also allows me to overcome the selection problem mentioned in Carozzi and Roth (2022) relative to the endogenous location of monitoring stations within urban areas[12]. The entity of data loss when using ground-based data is considerable: $PM_{2.5}$ data from the National Atmospheric Surveillance Program (NAPS) is only available for 61 DAs in 2000, growing to 230 in 2018 as new monitoring stations get added every year (see Figure A.1). Nonetheless, the satellite-retrieved measurements from Meng *et al.* (2019), when restricted to the DAs with at least one $PM_{2.5}$ ground monitoring station, correlate well with the NAPS readings, as shown in Figure A.2.

I rely on the Meng *et al.* (2019) $PM_{2.5}$ estimates in order to produce my main results. However, I also run the main analysis using $PM_{2.5}$ concentration data from van Donkelaar *et al.* (2019), as done e.g. in Sager and Singer (2022). While the two estimates are highly related, with a Pearson correlation coefficient of 0.795 (see Figure A.3), the concentrations from Meng *et al.* (2019) are generally lower throughout the sample. Moreover, a closer inspection of the van Donkelaar *et al.* (2019) rasters reveals that, beginning with the year 2004, much of the variability of $PM_{2.5}$ pixel values over Canadian CMAs is swept out, resulting in unrealistic estimates of pollution concentrations, especially with respect to their distribution over densely populated DAs[13].

Aware of a burgeoning literature relating population density and air pollution (Carozzi and Roth, 2022; Borck and Schrauth, 2021), I obtain population counts at the DA level from Rose *et al.* (2020), which are available for all years between 2000-2018[14].

[12]Monitoring stations are likely to be located where air pollution is higher, thereby introducing measurement error in an eventual empirical analysis.

[13]Moreover, the choice of employing data from Meng *et al.* (2019) is conservative, as results using the van Donkelaar *et al.* (2019) dataset are generally higher in magnitude.

[14]The dataset also contains population counts for all DAs extrapolated from Canadian censuses; however, this data is only available in 5-years intervals between 2001 and 2016.

Further, I employ the Canadian censuses to retrieve information on median income at the DA level, and extract the 2006 Material Deprivation Index from Pampalon *et al.* (2012) for all DAs in my sample, in order to inspect whether the tax has produced heterogeneous impacts along these dimensions. If the carbon tax was successful in producing a behavioural adjustment in BC residents, an expected result would be higher take up of alternative means of transport within metropolitan areas. Therefore, I leverage the detailed information contained in the four waves of Canadian census data between 2001-2016 to retrieve DA-level data on commute modes. I divide commute modes in two different categories: high emissions (cars, taxis, and motorcycles), and low emissions (public transport, bicycles, and walking)[15].

Figure 1 plots the baseline spatial distribution of the dependent variable and the main covariates over the Vancouver CMA, the most populated metropolitan area in the treated province of British Columbia. Time-varying variables are averaged over 2005-2007, the three years preceding the implementation of the carbon tax, while all variables retrieved from the Canadian Census are taken at their 2006 values, the last observation before the tax was instituted. The spatial distibution of $PM_{2.5}$ concentrations is as expected: values are indeed higher in central areas rather than in the periphery, qualitatively adhering to the traditional association with population density found e.g. in the US (Carozzi and Roth, 2022) or Germany (Borck and Schrauth, 2021); moreover, population density, material deprivation and the inverse of median income are all highly spatially correlated with air pollution at the baseline. Baseline commute mode seems to be inversely related with the spatial distribution of $PM_{2.5}$: areas whose inhabitants are less reliant on cars, taxis and motorbikes seem to be more polluted on average, a result probably due to their centrality with respect to the road networks and urban form. Summary statistics for the whole sample, split across treatment and control CMAs, are presented in Table A.1 and Table A.2 for the pre-treatment and post-treatment periods, respectively.

Lastly, I obtain monthly information on the BC gasoline and diesel fuel markets, at the province level, for January 1991-December 2016. In particular, I extract the annual sales of transportation fuels (motor gasoline and diesel), from Statistics Canada (2021b), gasoline and diesel price data from Kalibrate (formerly Kent Group Ltd.) at the monthly level for the city of Vancouver, which I consider representative of the entire province, monthly after tax income and unemployment rate data from Statistics

[15] I further decompose the low emissions category into public transport only and zero emissions commutes (cycling and walking).

Canada (2021c), and the CAD-USD monthly exchange rate, retrieved from the Pacific Exchange Rate Service at University of British Columbia's Sauder School of Business.

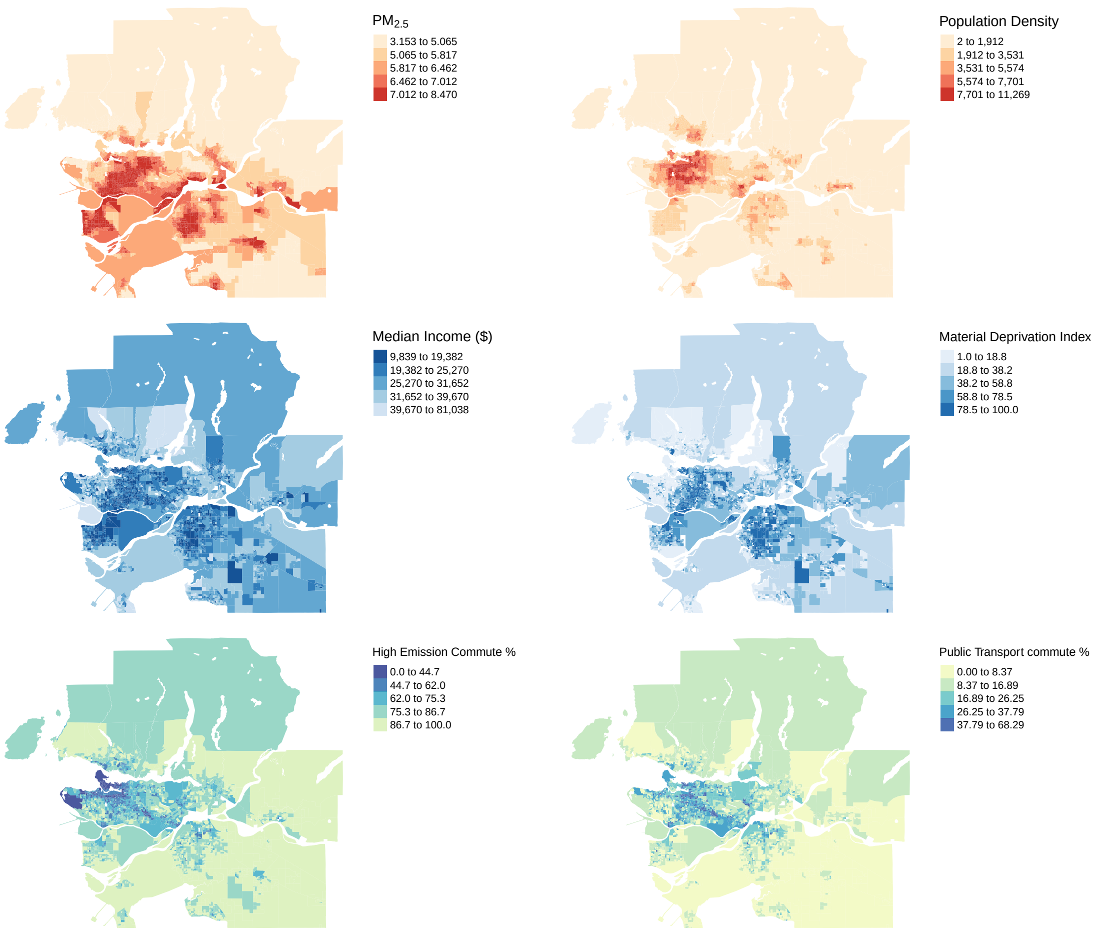

Figure 1: Spatial distribution of $PM_{2.5}$ and relevant covariates within the Vancouver CMA. Top row: $PM_{2.5}$ and population density; Middle row: median income and Material Deprivation Index; Bottom row: high emission commute mode % and public transport %.

3. Empirical Strategy

3.1. Two-way fixed effects difference-in-differences (TWFE-DID)

The core aim of my empirical strategy is to estimate the treatment effect of the 2008 British Columbian carbon tax on local air pollution, measured in terms of $PM_{2.5}$ concentrations at the DA level. A traditional methodology for this estimation is a two-way fixed effects difference-in-differences (TWFE-DID) regression. The estimating equation takes the form:

$$PM_{2.5it} = \beta TAX_{it} + \theta_t + \eta_i + \epsilon_{it} \tag{2}$$

Where TAX_{it} is the DID binary indicator, taking value 1 for all treated units after the implementation of the carbon tax in 2008, and 0 for all other observations; θ_t and η_i are respectively time and unit specific fixed effects, ϵ_{it} is a time-varying idiosyncratic error term, and β is the coefficient of interest, i.e. the average effect of being exposed to the carbon tax.

In order for β to be equal to the average treatment effect on the treated cohort (ATT), the identifying assumption is that parallel outcome trends between the treated and the control units hold, i.e. if the 2008 carbon tax had not been implemented in British Columbia, $PM_{2.5}$ levels in British Columbian DAs would have followed the same trajectory as $PM_{2.5}$ levels in DAs located in other Canadian provinces. Figure A.4 and Figure A.5 report the average $PM_{2.5}$ trends for 2000-2016 and 2000-2018, respectively, for British Columbian and control DAs. In both cases there is reason to suspect that a TWFE-DID regression would fail to identify the correct ATT: by giving equal weight to all control observations, TWFE-DID will indeed include units whose pre and post-treatment outcome paths fundamentally differ from those of DAs in British Columbia, and with greater potential for abatement[16].

It is also worth noting that in both cases, in addition to diverging trends, the level of $PM_{2.5}$ pollution is almost always[17] lower for British Columbian vis-à-vis control DAs. Province-specific factors such as city morphology, more progressive environmental attitudes, different car fleet compositions and heterogeneous availability of alternative means of transportation could be the reason why trends and levels diverge across British Columbian CMAs and control Provinces[18].

[16] An event study plot of differences in $PM_{2.5}$ between BC and control units' DAs (see Figure A.6) displays significant and unstable pre-treatment differences.

[17] Except for the van Donkelaar *et al.* (2019) dataset in the very first year of the panel, 2000.

[18] Treatment status in this instance is place-based and dependent on the political choice of an

3.2. Synthetic control method and synthetic difference-in-differences

A traditional solution to diverging pre-treatment trends in empirical applications (usually with a unique treated unit, but extensible to the case of multiple treated units) is the SCM (Abadie and Gardeazabal, 2003; Abadie, 2021). In the BC carbon tax case, the SCM constructs a set of synthetic DAs as a weighted combination of control DAs by finding, for each treated unit i, a non-negative vector of weights ω_i^{sc} summing to one, which ensures that each convex combination of the outcome variable for control units matches each outcome variable for the treated units for all periods up to the intervention date. Through this procedure, reliance on the parallel trends assumption, which is violated in my setting, is fundamentally weakened.

In order to combine the attractive features of both TWFE-DID (the inclusion of additive unit-specific and time-specific fixed effects), and SCM (reducing the reliance on the parallel trends assumption by weighting observations in order to ensure closely matched pre-intervention trends), Arkhangelsky *et al.* (2021) have introduced a new method, synthetic difference-in-differences (SDID), which employs time and unit (two-way) fixed effects in the regression function (as in TWFE-DID), together with unit-specific weights (as in SCM) and time-specific weights which lessen the role of time periods that are largely divergent from post-treatment time periods. In a nutshell, for each treated unit SDID estimates: (1) unit weights ω_i^{sdid} which underpin a synthetic control whose outcome is approximately parallel to the outcome for the treated unit; (2) time weights λ_t^{sdid} which ensure that the average post-treatment outcome for control units only differs by a constant from the weighted average of pre-treatment outcome for the control units – a synthetic pre-treatment period using controls. Once unit and time weights are calculated, SDID estimates a TWFE regression on the resulting panel, identifying the SDID ATT $\hat{\tau}^{sdid}$ by solving the minimisation problem:[19].

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\eta}, \hat{\theta}) = \underset{\tau,\mu,\eta,\theta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \eta_i - \theta_t - \tau TAX_{it})^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (3)$$

individual province, although sufficiently exogenous in timing (Rivers and Schaufele, 2015). Nonetheless, as in Sager and Singer (2022), bias in the TWFE-DID estimator introduced by the failure of the parallel trends assumption needs to be acknowledged and a different estimation strategy can give rise to more precise estimates.

[19] Section B presents a detailed formal comparison between TWFE-DID, SCM, and SDID, drawing on the seminal work of Arkhangelsky *et al.* (2021).

In the remainder of the paper, I regard SDID as my preferred method in order to estimate the effect of the 2008 BC carbon tax on air pollution co-benefits, as the methodology allows me to overcome the apparent violation of the parallel trends assumption in conventional TWFE-DID; nonetheless, I estimate my main regression and robustness checks using all three of TWFE-DID, SCM and SDID, in order to inspect the direction of the bias. I calculate standard errors for all methods using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021, p. 4109), with 200 replications. The procedure constructs a bootstrap dataset by sampling a portion of the original dataset with replacement, and computes the SDID estimator $\tau^{(b)}$ on this subset for each iteration b. The variance is then defined as:

$$\hat{V}_{\tau}^{b} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\tau}^{(b)} - \frac{1}{B}\sum_{b=1}^{B}\hat{\tau}^{(b)}\right)^2 \tag{4}$$

4. Results

In Figure 2 and Table 1, I report the results of the TWFE-DID, SCM and SDID regressions, using the Meng *et al.* (2019) $PM_{2.5}$ dataset. In the leftmost panel of Figure 2, it can be inferred how the baseline TWFE-DID strategy suffers from a violation of its foundational parallel trends assumption. The graphical representation of the regression analysis aids this line of interpretation: the DID ATT is indeed estimated by assuming that the outcome path of the treated units is parallel to the outcome path of the controls, thus the coefficient, $\hat{\tau}^{did} = 0.393$ is upward biased. In the centre panel of Figure 2, I plot the average outcome path for the treated units and the traditional synthetic control. The improvement in pre-treatment fit is dramatic, with a minimal average deviation between British Columbian DAs and their controls, implying that the SCM performs well in giving positive weights to control units which best approximate treated DAs' outcome paths and zero weight to control units which exhibit different trends. The direction of bias from the TWFE-DID regression is positive: SCM indeed identifies an effect of opposite sign to TWFE-DID, $\hat{\tau}^{sc} = -0.142$. Results for the SDID estimator are graphically shown in the right-most panel of Figure 2. At the bottom of the panel, pre-treatment time-weights are represented in pink. The estimator gives positive temporal weights to periods for which the treated and control units exhibit similar trends.

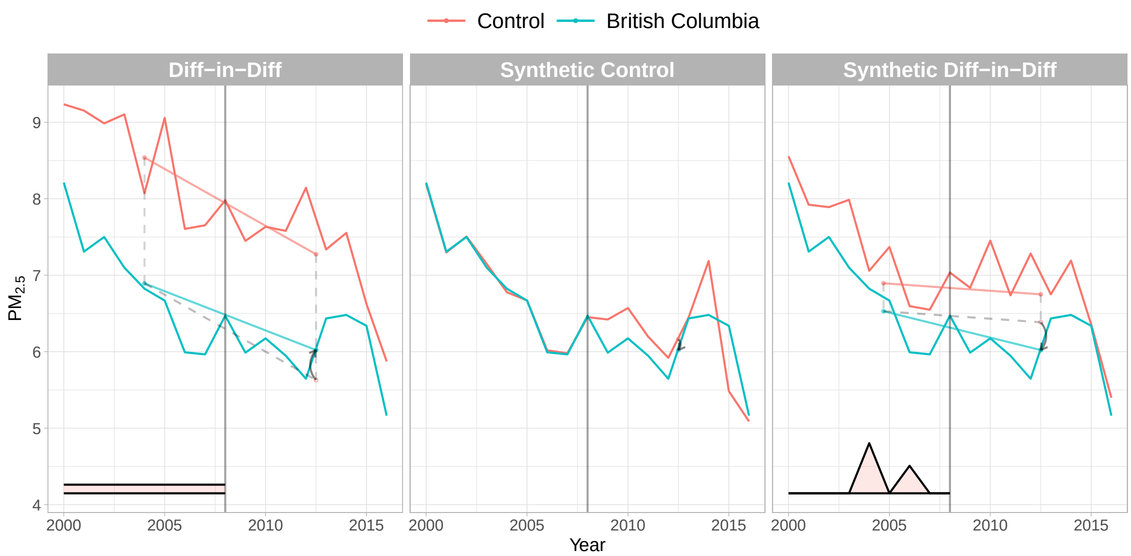

Figure 2: Graphical results from DID, SCM and SDID for $PM_{2.5}$ concentrations, with Meng *et al.* (2019) data. The 2008 carbon tax is denoted by a black vertical line.

The SDID estimator does a particularly good job in imposing pre-treatment parallel trends in the years preceding the tax, even if weights λ_t are unevenly distributed over the pre-intervention period. However, negligible weights in 2007-2008 are reassuring, given that a standard caveat in event-study methodologies is the excessive reliance on the single period immediately preceding the intervention (Heckman and Smith, 1999). The SDID procedure is able to select control units which exhibit pre-treatment trends that are almost perfectly parallel to BC's outcome path, especially in the four-year window preceding the intervention. The estimated ATT is $\hat{\tau}^{sdid} = -0.363$, higher than the SCM ATT, and corresponding to a 5.2% reduction with respect to pre-intervention mean pollution levels. I regard SDID as the preferred methodology due to its greater flexibility and to the selection of a sparser set of control DAs[20] While SCM obtains a near-perfect fit pre-treatment, the outcome path of its synthetic unit heavily depends on the particular set of units receiving positive weights, which in my highly disaggregated setting is not ideal[21].

Table 1: Summary of $\hat{\tau}$ point estimates and standard errors from all estimation methods, dependent variable from Meng *et al.* (2019).

	(1) ***DID***	(2) ***SCM***	(3) ***SDID***
$\hat{\boldsymbol{\tau}}$	0.3925	-0.1421	-0.3633
	(0.0074)	(0.0809)	(0.0219)
Unit FE	✓		✓
Year FE	✓	✓	✓
ω_i		✓	✓
λ_t			✓
$\boldsymbol{N_{obs}}$	432939	432939	432939

Notes: All point estimates represent the average impact of the 2008 carbon tax during the 2009-2016 post-treatment period. Standard errors are calculated using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021) with 200 replications. All regressions use 2000-2016 data.

[20]SDID selects indeed 6,258 control units among the untreated DAs and then performs DID on the matched sample with the inclusion of unit and time fixed effects to aid the estimation.

[21]In Figure D.1, I aggregate all 6,258 DAs which receive positive weights to the CMA level, in order to obtain the composition of synthetic BC in terms of percentages of other Canadian CMAs, in a similar vein to the traditional SCM methodology of Abadie (2021).

5. Mechanisms

5.1. Reductions in Transport Fuel Demand

The first candidate explanation for the observed reductions in particulate matter concentrations is a change in consumer behaviour regarding transportation fuel. Some evidence supporting this explanation is found in early analyses of the BC carbon tax (e.g. Rivers and Schaufele, 2015; Lawley and Thivierge, 2018), which use a limited post-intervention time period and only focus on gasoline consumption[22]. On the contrary, fuel substitution away from gasoline and towards diesel is claimed to be a potential mechanism behind the $PM_{2.5}$ increases found in Saberian (2017), notwithstanding the negative impacts found by the time series analysis of Bernard and Kichian (2019) and the strong prevalence of gasoline vehicles among BC car sales (see Figure C.3 and C.4).

I reconcile the evidence on the aggregate level effects of the carbon tax on transportation fuel demand by introducing a recently developed method for high-frequency time series analysis: the Causal-ARIMA (C-ARIMA) estimator of Menchetti *et al.* (2022). By exploiting features of ARIMA models, the method is especially appropriate to analyse complex seasonal, nonstationary processes such as gasoline and diesel sales observed monthly (see Figure C.1, panels A and B). C-ARIMA combines attractive features from the DID and SCM estimator for the case in which no suitable control unit is available[23] and when the number of pre-intervention time periods is large[24]. Under standard assumptions[25], C-ARIMA is able to learn the treated unit's time series dynamics and forecast it after the shock takes place. By using the forecasted series as the treated unit's counterfactual outcome, the method identifies two main sets of causal effects: the temporal average causal effect and the cumulative treatment effect.

I run C-ARIMA separately for per capita monthly gasoline and diesel sales at the aggregate BC level between January 1991 and December 2016. The intervention date is July 2008, i.e. the specific month in which the BC carbon tax came into effect. In Table 2, I report the results from estimations with and without a matrix

[22]Which accounts for most of the residential vehicle fleet (see Figure C.3) but does not include heavy duty vehicles used in commercial and industrial operations (Bernard and Kichian, 2019).

[23]In my context, a pool of eligible control units is represented by other Canadian provinces. However, other provinces exhibit diverging pre-intervention trends in gasoline sales (see Figure C.2) when aggregating the TWFE-DID coefficients into an event study plot.

[24]As is the case in the monthly analysis of BC fuel consumption between January 1991 and December 2016, with 210 pre-intervention time periods.

[25]No temporal interference (i.e. absence of anticipation effects), covariates-treatment independence and conditional stationarity (Menchetti *et al.*, 2022).

of business cycle controls[26]. Both the temporal average causal effect $\hat{\tau}_t$ and the cumulative causal effect $\sum_{t=t_{int}}^{T} \hat{\tau}_t$ are negative and statistically significant across all specifications, highlighting a successful impact of the BC carbon tax in decreasing fuel demand, consistently with Rivers and Schaufele (2015) and Bernard and Kichian (2019). In Figure C.1, the results from the estimation are reported graphically.

Table 2: C-ARIMA: Monthly Gasoline and Diesel Demand

	Gasoline Sales		Diesel Sales	
	(1)	(2)	(3)	(4)
$\hat{\tau}_t$	-3.883***	-4.675***	-1.756***	-0.912***
	(0.553)	(0.506)	(0.412)	(0.236)
$\sum_{t=t_{int}}^{T} \hat{\tau}_t$	-396.052***	-818.405***	-179.089***	-92.983***
	(56.453)	(14.962)	(42.066)	(24.093)
Controls	-	✓	-	✓
Observations	312	312	312	312

Notes: The dependent variable is total monthly gasoline (diesel) sales per capita (in litres) recorded in British Columbia between January 1991 and December 2016. $\hat{\tau}_t$ denotes the Temporal Average Causal Effect. $\sum_{t=t_{int}}^{T} \hat{\tau}_t$ denotes the Cumulative Causal Effect from the intervention period t_{int} to the last period under consideration T. Columns (2) and (4) include a matrix of monthly province-level covariates, namely consumer price index, gasoline (diesel) crude cost, population, unemployment rate, after tax income and the US-CAD exchange rate. Significance levels ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$

5.2. Commute Mode Switching

I analyse commute mode choices at the DA level as an additional mechanism driving the main results. While commute mode is an imperfect measure of the number and type of trips made by British Columbians, I can rely on the same administrative level to the one used in the main analysis by retrieving information from the 2001, 2006, 2011, and 2016 Canadian censuses, thereby preserving granularity. In Table 3 and Table 4, I report TWFE-DID regression results[27] employing the share of

[26] Namely, provincial population, unemployment, after tax income, exchange rate and the cost of crude gasoline and diesel, respectively.

[27] Due to the structure of the data, collected at 5-year intervals, I am prevented from using the SCM and SDID methodology in this exercise; I thus resort to traditional TWFE-DID estimation of commute mode switching, analysing the data separately for each category of commute mode. Details on this estimation strategy are reported in Section C.2.

commuters using high-emissions and public transport commute modes, respectively [28].

In all tables, column (1) is the baseline specification, a simple TWFE-DID regression with DA and year fixed effects and no controls, employing the full panel of DAs across census years. In column (4), I add weather controls for precipitation, maximum and minimum temperature, and wind speed, plus the natural logarithm of population and median income. When employing the full pool of control DAs, the first result of note is that British Columbian DAs experience an average 4.2% reduction in the use of cars, taxis, and motorcycles, which rises to 4.7% when adding controls. This reduction is almost specular to the increase in the share of commuters using public transport, biking and walking to reach their workplace (Table C.1). Moreover, as evidenced in Table 4 and Table C.2, most of this increase (3.5-3.9%) is due to a higher reliance on public transport, while a residual share of 0.5-0.7% is due to a switch to active commuting. All results are confirmed and stronger in magnitude when considering more restrictive specifications: columns (2) and (5) restrict the specifications in (1) and (4) to the DAs which receive positive weights in the main SDID regressions, in order to establish whether the mechanisms are effectively retrieved when employing the same set of observations on which the main ATT is estimated. Results are higher in magnitude by about 1%, jumping to a 5.3% reduction in high-emission commute modes in the case without controls. Here, the inclusion of control variables slightly dampens the impact to 5.2%; nonetheless, the specularity with the increase in low-emission commute modes is preserved. Finally, in columns (3) and (6) I augment the TWFE-DID regressions by retrieving an including the weights from the main SDID regressions. I weigh all treatment observations equally and all control observations according to the value of ω_i they receive after the data-driven SDID procedure. The magnitude of the increase in public transport commute share increases further, to 4.2% in the case without covariates and is again dampened to 4.1% in the case with covariates. The hypothesis of a behavioural adjustment by BC citizens in response to the carbon tax is thus confirmed; residents of BC's DAs switch away from high-emissions commute modes towards low-emissions ones, with public transport as the main container for these substitutions.

[28]As the low-emissions transport mode is the sum of public transport and zero-emissions modes, I only report the results for public transport in the main text and present the aggregate low emissions and the sub-split for zero-emissions in Table C.1, and Table C.2.

Table 3: TWFE-DID results for high emissions commute mode

	High Emission Commute Mode					
	(1)	(2)	(3)	(4)	(5)	(6)
DID	-0.0417***	-0.0527***	-0.0549***	-0.0466***	-0.0519***	-0.0516***
	(0.0105)	(0.0095)	(0.0103)	(0.0102)	(0.0106)	(0.0109)
DA FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls				✓	✓	✓
SDID control pool		✓	✓		✓	✓
SDID weights			✓			✓
R^2	0.87184	0.83989	0.84360	0.87595	0.84508	0.84847
Adjusted R^2	0.82896	0.78629	0.79124	0.83400	0.79267	0.79721
Observations	101,358	38,769	38,769	100,244	38,348	38,348

Notes: The dependent variable is the dissemination area level share of high emissions commutes. All regressions include dissemination area and year fixed effects. Columns (4)-(6) include controls for precipitation, maximum and minimum temperature, and wind speed, plus the natural logarithm of population and median income. Columns (2), (3), (5) and (6) restrict the control unit pool to DAs which receive positive weights in the main SDID regression. Columns (3) and (6) additionally include the estimated SDID unit weights ω_i as regression weights. Standard errors are clustered at the CMA level. ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$

Table 4: TWFE-DID results for public transport

	Public Transport Commute Mode					
	(1)	(2)	(3)	(4)	(5)	(6)
DID	0.0352***	0.0410***	0.0417***	0.0391***	0.0422***	0.0414***
	(0.0107)	(0.0107)	(0.0112)	(0.0115)	(0.0115)	(0.0111)
DA FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls				✓	✓	✓
SDID control pool		✓	✓		✓	✓
SDID weights			✓			✓
R^2	0.83768	0.78668	0.78011	0.84196	0.79197	0.78571
Adjusted R^2	0.78336	0.71526	0.70650	0.78851	0.72160	0.71322
Observations	101,358	38,769	38,769	100,244	38,348	38,348

Notes: The dependent variable is the dissemination area level share of public transport commutes. All regressions include dissemination area and year fixed effects. Columns (4)-(6) include controls for precipitation, maximum and minimum temperature, and wind speed, plus the natural logarithm of population and median income. Columns (2), (3), (5) and (6) restrict the control unit pool to DAs which receive positive weights in the main SDID regression. Columns (3) and (6) additionally include the estimated SDID unit weights ω_i as regression weights. Standard errors are clustered at the CMA level. ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$

6. Robustness Checks

In this section, I test the consistency of the main results by employing the van Donkelaar *et al.* (2019) $PM_{2.5}$ dataset, and performing additional analyses on subsamples of the full dataset. As I show in the following sections, my results are robust in each specification.

6.1. Main results with van Donkelaar *et al.* (2019) $PM_{2.5}$ data

I repeat the TWFE-DID, SCM and SDID estimation using the van Donkelaar *et al.* (2019) $PM_{2.5}$ dataset, which is available between 2000 and 2018. Notwithstanding the high correlation between the two outcome variables, as outlined in Figure A.3, both the treatment and control pre-intervention trends exhibit some differences with respect to the Meng *et al.* (2019) dataset[29]. The violation of the parallel trends assumption is once again highlighted in the graphical representation of the TWFE-DID regression in Figure D.3, which, differently from the previous estimation, identifies a negative effect of the 2008 carbon tax on emissions of $\hat{\tau}^{did} = -0.495$ (see Table D.1).

The SCM, represented graphically in the middle panel of Figure D.3, again obtains a good pre-treatment fit, signalling that each British Columbian DA's outcome path is best approximated by a convex combination of control DAs rather than equally weighted control units. Furthermore, as evidenced in Table D.1, the direction of the TWFE-DID bias is confirmed: the SCM estimates a negative ATT of $\hat{\tau}^{sc} = -0.709$, therefore qualitatively reinforcing the SCM result of Table 1. A similar conclusion can be drawn from the results of the SDID estimation, presented in the right-most panel of Figure D.3. The SDID procedure is able to select control units[30] which exhibit pre-treatment trends that are almost perfectly parallel to BC's outcome path, with the exception of outlying time periods which receive zero-weights in the estimation. The estimated ATT is $\hat{\tau}^{sdid} = -0.890$, therefore slightly lower, but qualitatively similar to the SCM ATT. In terms of magnitude, both the SCM and SDID regressions identify a substantial drop in $PM_{2.5}$ concentrations with respect to 2000-2007 levels, corresponding to a reduction of 10.9% from the pre-intervention $PM_{2.5}$ mean for British Columbia.

[29] However, the temporal location of peaks and troughs is generally respected, as is the relationship between the BC and control units outcome path. Indeed, DAs located in British Columbia always exhibit lower average annual concentrations of particulate pollution, and their $PM_{2.5}$ trend prior to 2008 appears to decline at an even faster pace than for control observations, barring some peaks in concentrations typical of the control provinces.

[30] The composition of the donor pool, aggregated to the CMA level, is reported in Figure D.2.

6.2. Narrower Spatial and Temporal Scope

First, I restrict the treated pool to DAs within the Vancouver metropolitan area, excluding all DAs in the Abbotsford and Victoria CMAs. The resulting treatment cohort is comprised of 2874 DAs, vis-à-vis the 3490 DAs constituting the entire treatment unit pool; the control pool is kept the same, with 21989 control DAs. I then run TWFE-DID, SCM and SDID on the restricted sample, in order to be able to compare my results with those obtained by Saberian (2017). Perhaps unsurprisingly, given the relatively small number of DAs pertaining to the Abbotsford and Victoria CMAs, the results (reported in Figure D.4 and Table D.2) are qualitatively unchanged from the main regressions using the Meng *et al.* (2019) dataset.

Secondly, I restrict the dataset to those DAs corresponding to the location of NAPS monitoring stations (see Figure A.1), by spatially joining monitoring stations' locations to DAs[31]. Here, the size of the dataset is considerably restricted: the cross-section of DAs kept in the treated pool counts just 25 observations, while 106 DAs are kept in the control pool. This exercise allows me to infer whether my results also arise when considering just those locations in which pollution monitors have been established, thereby restricting the analysis to areas in which pollution is likely to be a greater concern. Once again, the results (presented in Figure D.5 and Table D.3), are qualitatively similar to the main specifications, a first signal that air pollution experiences greater reductions in places which exhibit greater levels of pre-intervention concerns about air quality. Notably, the performance of the SDID estimator is not considerably worsened on this much smaller sample, with both estimators achieving a reasonable pre-treatment fit, and therefore identifying credible ATTs. On the contrary, the fit of the SCM seems to be substantially worse, and the method identifies a much higher ATT than in other specifications.

Lastly, I restrict the estimation window to 2000-2013, for two main reasons: checking whether the carbon tax ramp-up is the main mechanism behind the continuous reductions [32], and comparing my results with Saberian (2017). The results, presented in Figure D.6 and Table D.4 identify a much higher ATT of $\hat{\tau}^{sdid} = -0.67$, which corroborates the first hypothesis and is not comparable with the study by Saberian (2017), which identified an increase in particulate pollution over the same temporal window, possibly due to selection bias in the establishment of monitoring stations.

[31] I match DAs with all monitoring stations in the dataset, regardless of the date of establishment of each monitoring station, in order to maximise observations.

[32] In 2012, the carbon tax was frozen at \$30/tCO_2 as reported in Section 1.

6.3. Using NAPS Monitoring Stations

A further analysis is reported in Table D.5 and D.6. Here, I depart from using remotely sensed measurements, employing monthly mean concentrations and the number exceedances of the daily safe threshold of $PM_{2.5}$[33] retrieved from NAPS monitoring stations. As in Saberian (2017), all analyses employ the TWFE-DID estimator with unit and month-year fixed effects. The results obtained by Saberian (2017) are confirmed for a panel of the top 15 Canadian cities for 1998-2013 (Table D.6) for both concentrations and exceedances. However, the detrimental effect of the carbon tax on air quality is not statistically significant in any specification exploiting the full panel of NAPS stations, and when extending the panel to 2016 (Table D.5). As shown in Section 4, the TWFE-DID estimator exhibits positive bias in this setting, which is the likely explanation for this puzzle.

[33] The daily safe threshold for $PM_{2.5}$ is 25 $\mu g/m^3$.

7. Environmental Justice Gaps

In light of a growing literature in environmental justice (e.g. Sager and Singer, 2022; Hernandez-Cortes and Meng, 2023; Grainger and Ruangmas, 2018), I examine efficiency-equity trade-offs in the realisation of co-benefits, analysing whether the estimated air pollution reductions arise heterogeneously over metropolitan areas. In the main analysis, the parameter identifying the effect of the 2008 BC carbon tax on $PM_{2.5}$ emissions has always been assumed as constant across treated units. Nonetheless, when dealing with disaggregated data within Census Metropolitan Areas, a homogeneously estimated ATT is likely to mask substantial heterogeneities across DAs which could be highly informative about the performance of different locations within metropolitan areas.

A first channel to examine is certainly that of pre-existing pollution levels: standard economic theory would in fact predict that emission abatement would happen first where the marginal cost of reducing emissions is lower, i.e. where pre-existing pollution is higher (that is, lower-hanging fruits would be picked earlier). This avenue is explored by Sager and Singer (2022) and Auffhammer *et al.* (2009), who find substantially higher reductions in $PM_{2.5}$ and PM_{10} due to the Clean Air Act in nonattainment US census tracts that are more polluted in the three years preceding the implementation of the policy. In light of the results of Carozzi and Roth (2022) and Borck and Schrauth (2021), it is also worth exploring whether heterogeneity in air pollution reductions arises at different levels of the population density distribution: indeed, while densely populated areas have been shown to experience higher concentrations of $PM_{2.5}$ particulate, usually population density is higher in city centres, where greater opportunities for substitution away from cars may arise. Lastly, an unexplored channel in carbon pricing is that of "spatial regressiveness". A large body of research has shown that carbon pricing is regressive along income and wealth dimensions, but the relationship between the geographic distribution of income and wealth and the burden of carbon taxation is relatively underinvestigated. Therefore, I also inspect the heterogeneity of $PM_{2.5}$ reductions with respect to the geographic distribution of median income and the material deprivation index.

As the SDID methodology does not allow the inclusion of interactions in the estimation procedure, I split the treatment sample into quintiles of baseline[34] $PM_{2.5}$, population

[34]For time-varying covariates I use the average of the three years prior to treatment as the baseline value; for variables retrieved from the Canadian census, I use their 2006 values, i.e. the last observation prior to the implementation of the carbon tax.

density, median income and material deprivation index. I then run SDID separately for each quintile and, in Figure 3, I summarise the results graphically.

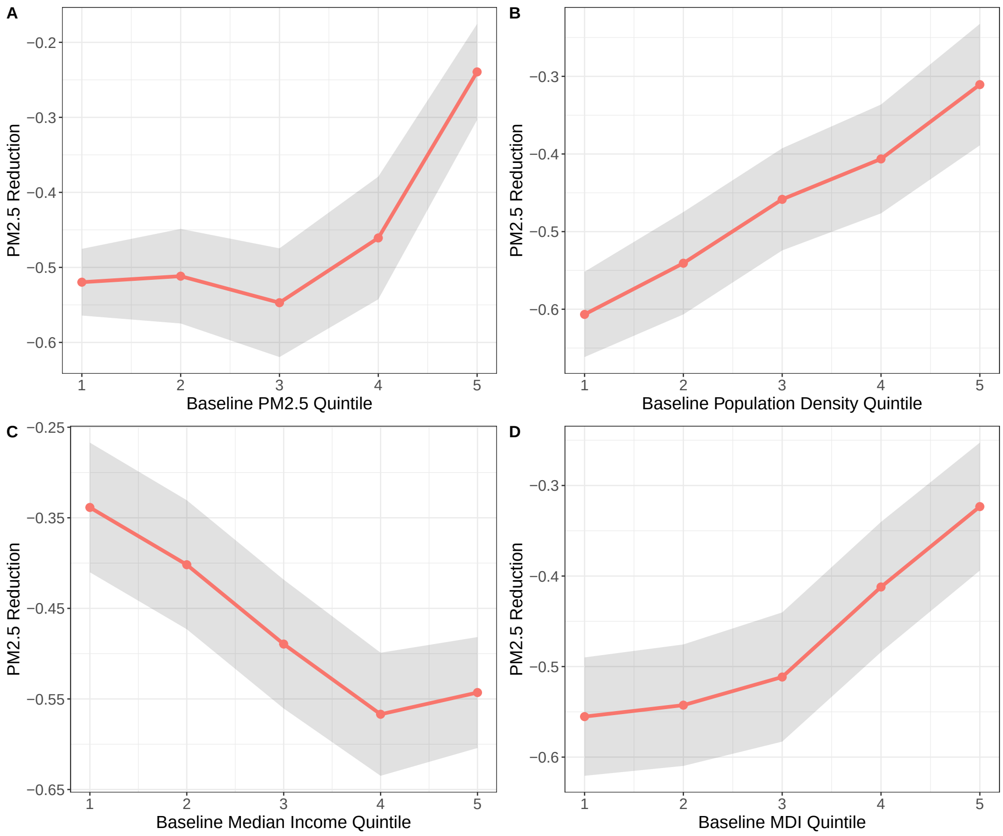

Figure 3: Graphical SDID results: heterogeneity by quintiles of baseline characteristics for the treated sample. Panel A) Quintiles of baseline $PM_{2.5}$; B) Quintiles of baseline population density; C) Quintiles of baseline median income; D) Quintiles of baseline material deprivation index. ATT point estimates reported in red, with 95% confidence intervals shaded in grey.

Quintile-SDID results for baseline $PM_{2.5}$ concentrations are presented in panel A of Figure 3. It is immediate to infer that greater reductions arise in DAs with lower pollution levels between 2005 and 2007. The ATT is in the [-0.2,-0.6] range, with the lowest quintiles of baseline pollution experiencing 2.2 times larger reductions with respect to the highest quintiles. Panel (B), which shows the SDID effects for quintiles of baseline population density, is consistent with the results for baseline pollution levels. Denser locations within metropolitan areas see lower reductions of particulate matter with respect to less dense DAs, underpinning a worsening of the pollution-density gap (Carozzi and Roth, 2022). Taken in conjunction, these insights appear to confirm that the 2008 carbon tax was not effective in curtailing traffic in more central areas within British Columbian metropolitan areas, but rather had greater effect in peri-urban locations. More surprising are the results in panel (C) and (D), which highlight the fact that relatively better off DAs within metropolitan areas have experienced greater reductions, possibly reflecting an inverse relationship between density and income, but more importantly signalling that the pollution-income gap has increased as a result of the carbon tax. This result is a clear confirmation of the "spatial regressiveness" hypothesis, i.e. that a carbon tax is not only regressive on the vertical income dimension, but also geographically, with greater gains in better off areas.

8. Health Gains

In order to understand the magnitude of the economic co-benefits from air pollution reductions arising due to the 2008 carbon tax, I convert the $PM_{2.5}$ estimates from Section 4 into a monetary quantification of the associated health gains. Notwithstanding the relatively low concentrations of particle pollution in the the British Columbian context, where pre-treatment air quality was of substantial better quality than in other North American locations (e.g. in the USA), it is important to note that the concept of "safe" thresholds for particle pollution concentrations is more normative than positive. Indeed, some studies (e.g. Krewski *et al.*, 2009) have highlighted that the marginal benefits from abatement may be nonlinear in baseline concentrations, with lower gains from abatement at higher levels of baseline air pollution. Hence, any improvement in air quality is likely to carry significant benefits in terms of reductions in mortality rates; moreover, the estimates reported in this section are a lower bound of the gains from local pollution reductions, as $PM_{2.5}$ has been shown to have a multidimensional impact, ranging from health to productivity, to cognition and the formation of human capital (Aguilar-Gomez *et al.*, 2022).

Drawing from Fowlie *et al.* (2019) and Carozzi and Roth (2022), my approach consists of two steps. I first estimate the impact of a reduction in $PM_{2.5}$ concentrations in terms of mortality reductions, using concentration-response ("hazard") functions derived from the environmental health literature. Second, I retrieve the central estimate of the willingness to pay (WTP) to avoid a premature death from Health Canada (2021) and Chestnut and De Civita (2009)[35], and multiply the mortality reductions estimated in the first step by the central estimate of the Value of a

[35]It must be noted that the reported estimate for the Value of a Statistical Life does not reflect directly the economic value of an individually identified person's life, but rather the aggregation of estimates of the WTP for a small reduction in mortality risk. Using the VSL central estimate of $6,500,000, for example, the average Canadian would be willing to pay $65 to reduce the risk of premature death by 1 out of 100,000.

Statistical Life (VSL), equal to \$6.5 million in 2007 Canadian dollars, for each DA in the census metropolitan areas of Vancouver, Victoria, and Abbotsford.

The traditional form of the Cox proportional hazard model used in the environmental health literature is the log-linear regression reported in Fowlie *et al.* (2019):

$$ln(\gamma) = \zeta + \alpha PM_{2.5} \tag{5}$$

Where $ln(\delta)$ is the natural logarithm of mortality risk, $\zeta = ln(Z)$, and $PM_{2.5}$ are the local pollution concentrations. The term Z is a vector of covariates other than $\mathrm{PM}_{2.5}$ which impact mortality, and can be rewritten as $Z = Z_0 + exp(\beta_1 x_1 + ... + \beta_n x_n)$, with Z_0 being the mortality risk when all covariates are zero. Indicating γ_0 as the baseline mortality risk, and rearranging terms[36], the change in mortality rate $\Delta\gamma$ can be related to the change in pollution levels $\Delta PM_{2.5}$ with the following equation:

$$\Delta\gamma = \gamma_0 \left(1 - \frac{1}{e^{\alpha \Delta PM_{2.5}}}\right) \tag{6}$$

In order to find the total number of deaths for each DA associated with the above change in mortality rate $\Delta\gamma$, this quantity needs to be multiplied by the population of each DA[37]:

$$\Delta Deaths_i = Population_i \left[\gamma_0 \left(1 - \frac{1}{e^{\alpha \Delta PM_{2.5}}}\right)\right] \tag{7}$$

And finally, the monetary health gains in terms of mortality reductions at the DA level, ΔY_i, are obtained by multiplying the above estimates by the VSL figure of \$6.5 million CAD obtained from Health Canada (2021):

$$\Delta Y_i = VSL * \Delta Deaths_i \tag{8}$$

Hence, in order to estimate the model outlined in Equation 6, and thus obtain mortality rate changes at the DA level, I first need to estimate the baseline mortality rate γ_0. Consistently with the literature, I obtain data for deaths due to lung cancers, all circulatory diseases, and all respiratory diseases from the ICD.10 selected causes of death at the CMA level from Statistics Canada (2021a). I divide total

[36] The derivation is as follows (Carozzi and Roth, 2022):

$$\Delta\gamma = Z(e^{\alpha PM^0_{2.5}} - e^{\alpha PM^1_{2.5}}) \rightarrow \Delta\gamma = Ze^{\alpha PM^0_{2.5}} \left[1 - e^{-\alpha(PM^0_{2.5} - PM^1_{2.5})}\right]$$

[37] I use the baseline population level, that is, the population of each DA in the year 2008.

deaths due to the listed causes by total CMA population, and assign the resulting (baseline) mortality rates to all DAs in a given CMA. The parameter α is usually not directly indicated in epidemiology studies, which instead report the relative risk (RR) increase due to a given increase in $PM_{2.5}$. For instance, Lepeule *et al.* (2012) report an all-cause RR of 1.14 associated with a $\Delta PM_{2.5}$ of 10 $\mu g/m^3$, while Krewski *et al.* (2009)'s estimate of RR is 1.06. However, it is straightforward to retrieve α by exploiting the relationship between RR and $\Delta PM_{2.5}$, as reported in Carozzi and Roth (2022): $\alpha = ln(RR)/\Delta PM_{2.5}$.

I employ these two estimates, in combination with the estimated $PM_{2.5}$ reductions for each quintile of the pre-intervention $PM_{2.5}$ distribution, in order to calculate the gains from mortality reductions at the DA level for the three CMAs included in the treated sample: Vancouver, Victoria and Abbotsford. In Figure 4, I visually report the results of this exercise for each CMA, using RR = 1.14 as estimated by Lepeule *et al.* (2012) (visual results using the RR estimate from Krewski *et al.* (2009) are reported in Figure F.1).

The left panel maps the estimated mortality reductions per 1000 people (estimated according to Equation 7), while the right panel shows the associated per capita health gains, estimated via Equation 8. The median per capita monetary gains due to the estimated reductions in $PM_{2.5}$ are large: \$198 when using the Lepeule *et al.* (2012) RR and \$88 with the RR from Krewski *et al.* (2009)[38].

The monetary value of per capita air quality co-benefits from the BC carbon tax is 1.7 times of the per capita low income climate action tax credit, i.e. the carbon tax rebate for low-income families[39]. Moreover, the total monetary value of co-benefits ranges between \$507.2 million and \$1.03 billion annually, or 40-81% of annual carbon tax revenues once the tax reached its \$30/tCO_2 level in 2012 (Ministry of Finance, 2013). The spatial distribution of these gains shows substantial heterogeneity: in particular, it is once again striking how air pollution co-benefits seem to be concentrated in peri-urban areas and positively correlated with income (see also Figure F.2). The results confirms that carbon taxation appears to be spatially regressive over urban areas, with greater co-benefits arising in higher income, low pollution tracts, underpinning increasing environmental justice gaps, as also evidenced in Section 7.

[38] The same gains are \$402 and \$178, respectively, if calculated using the ATT estimated with the van Donkelaar *et al.* (2019) $PM_{2.5}$ dataset instead of Meng *et al.* (2019).

[39] For this comparison, I use the last revision of the low income climate action tax credit, amounting to \$115.50 per adult plus \$34.50 per child (Ministry of Finance, 2013).

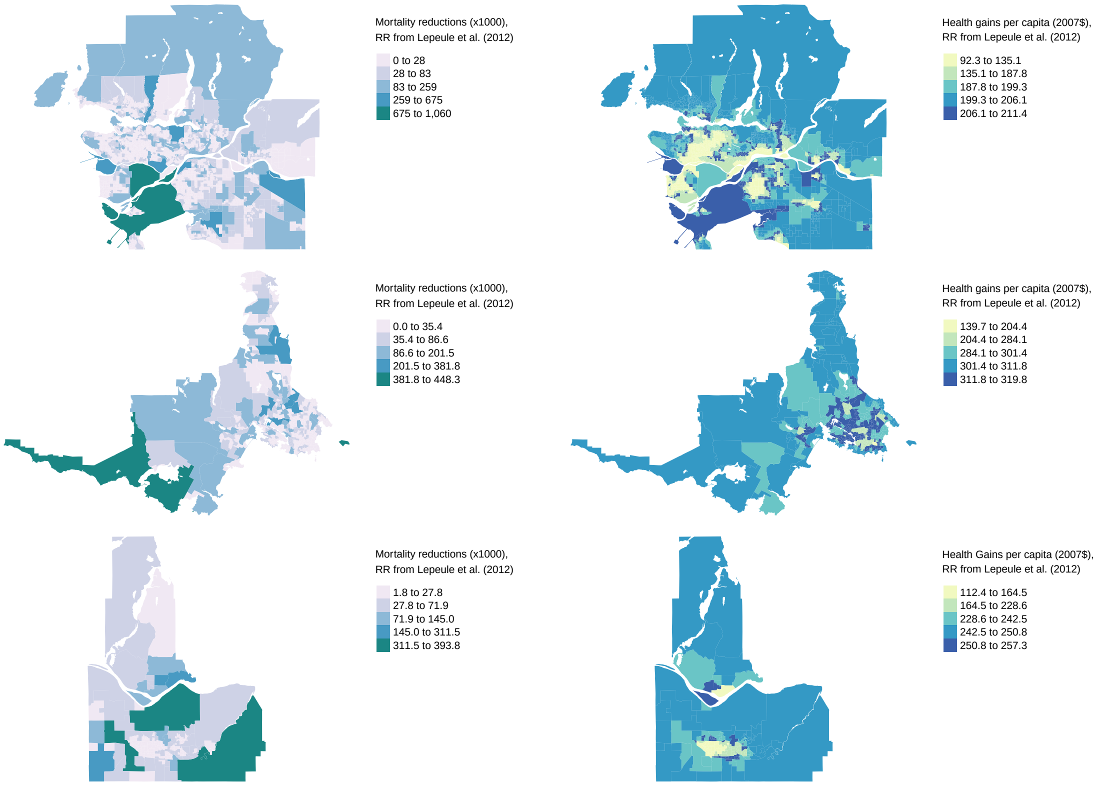

Figure 4: Spatial distribution of mortality reductions per 1000 residents (left panel) and health gains per capita (right panel) using the RR estimates from Lepeule *et al.* (2012), for the Vancouver (top row), Victoria (middle row) and Abbotsford (bottom row) CMAs.

9. Conclusions

This paper connects two areas of extreme concern for what regards environmental policy. Air pollution co-benefits from carbon taxation are likely to be large in magnitude and may partially or fully offset the costs of climate mitigation. Incorporating the monetary value of realised air quality improvements in cost-benefit analyses of carbon taxes is essential in order to correctly calibrate them and enhance their attractiveness. Conversely, environmental justice implications of market-based instruments are an often overlooked dimension due to the focus on efficiency, rather than equity. Ignoring potentially regressive consequences in terms of the societal distribution of co-benefits could hinder public support towards climate policy.

I show that the introduction of carbon pricing can significantly improve local air quality. After the implementation of the 2008 carbon tax, $PM_{2.5}$ concentrations dropped by 5.2-10.9% in British Columbian dissemination areas, compared to a counterfactual obtained through the synthetic difference-in-differences estimator. The air quality improvement is driven by reductions in fuel demand and by transport mode switching, mostly in favour of public transport. In terms of environmental justice, the estimated reductions are significantly heterogeneous across the geography of British Columbian census metropolitan areas, with greater effects found in less polluted, less dense areas and in better off neighbourhoods. These results highlight a spatial dimension of the regressive nature of carbon pricing: a carbon tax can indeed exacerbate the pre-existing pollution gap, and the pollution-income gap. Instruments designed to attenuate inequitable effects may then be designed in advance of the deployment of carbon pricing in order to smooth potentially regressive consequences.

Finally, I convert the improvements in air quality into reductions in mortality rates and monetary health gains from co-benefits of carbon taxation. With a median estimate of $198 per capita, the health gains are large and comparable to the rebates offered to low-income families in British Columbia to mitigate the impact of the tax on their disposable income. Health benefits are heterogeneously distributed across metropolitan areas and accrue primarily to neighbourhoods in higher income brackets, once again highlighting the need for redistribution in the design of climate policy.

References

Abadie, Alberto (2021). "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects". *Journal of Economic Literature* 59.2, pp. 391–425.

Abadie, Alberto, Diamond, Alexis, and Hainmueller, Jens (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program". *Journal of the American Statistical Association* 105.490, pp. 493–505.

Abadie, Alberto and Gardeazabal, Javier (2003). "The Economic Costs of Conflict : A Case Study of the Basque Country". *American Economic Review* 93.1, pp. 113–132.

Abatzoglou, John, Dobrowski, Solomon, Parks, Sean, and Hegewisch, Katherine (2018). "TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015". *Nature Scientific Data* 5, p. 170191.

Aguilar-Gomez, Sandra, Dwyer, Holt, Graff Zivin, Joshua S, and Neidell, Matthew J. (2022). *This is Air: The "Non-Health" Effects of Air Pollution.* Working Paper 29848. National Bureau of Economic Research.

Ahmadi, Younes, Yamazaki, Akio, and Kabore, Philippe (2022). "How Do Carbon Taxes Affect Emissions? Plant-Level Evidence from Manufacturing". *Environmental and Resource Economics* 82.2, pp. 285–325.

Andersson, Julius J. (2019). "Carbon Taxes and CO2 Emissions: Sweden as a Case Study". *American Economic Journal: Economic Policy* 11.4, pp. 1–30.

Arkhangelsky, Dmitry, Athey, Susan, Hirshberg, David A., Imbens, Guido W., and Wager, Stefan (2021). "Synthetic Difference-in-Differences". *American Economic Review* 111.12, pp. 4088–4118.

Auffhammer, Maximilian, Bento, Antonio, and Lowe, Scott (2009). "Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of a spatially disaggregated analysis". *Journal of Environmental Economics and Management* 58, pp. 15–26.

Azevedo, Deven, Wolff, Hendrik, and Yamazaki, Akio (2023). "Do carbon taxes kill jobs? Firm-level evidence from British Columbia". *Climate Change Economics* 14.02, pp. 1–46.

Basaglia, Piero, Behr, Sophie M., and Drupp, Moritz A. (2023). *De-Fueling Externalities: Causal Effects of Fuel Taxation and Mediating Mechanisms for Delivering Climate and Health Benefits.* CESifo Working Paper 10508.

Bernard, Jean-Thomas and Kichian, Maral (2019). "The long and short run effects of British Columbia's carbon tax on diesel demand". *Energy Policy* 131, pp. 380–389.

Bondy, Malvina, Roth, Sefi, and Sager, Lutz (2020). "Crime Is in the Air: The Contemporaneous Relationship between Air Pollution and Crime". *Journal of the Association of Environmental and Resource Economists* 7.3, pp. 555–585.

Borck, Rainald and Schrauth, Philipp (Jan. 2021). "Population density and urban air quality". *Regional Science and Urban Economics* 86, p. 103596.

Boyce, James and Pastor, Manuel (2013). "Clearing the air: incorporating air quality and environmental justice into climate policy". *Climatic Change* 120.4, pp. 801–814.

Carattini, Stefano, Baranzini, Andrea, Thalmann, Philippe, Varone, Frédéric, and Vöhringer, Frank (2017). "Green Taxes in a Post-Paris World: Are Millions of Nays Inevitable?" *Environmental and Resource Economics* 68.1, pp. 97–128.

Carattini, Stefano, Kallbekken, Steffen, and Orlov, Anton (2019). "How to win public support for a global carbon tax". *Nature* 565.

Carozzi, Felipe and Roth, Sefi (2022). "Dirty density: Air quality and the density of American cities". *Journal of Environmental Economics and Management*, p. 102767.

Chestnut, L.G. and De Civita, P. (2009). "Economic valuation of mortality risk reduction : review and recommendations for policy and regulatory analysis". *Government of Canada. Policy Research Initiative.*

Currie, Janet, Voorheis, John, and Walker, Reed (2020). *What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality.* en. Tech. rep. w26659. Cambridge, MA: National Bureau of Economic Research, w26659. (Visited on 08/24/2022).

Doudchenko, Nikolay and Imbens, Guido W (2016). *Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis.* Working Paper 22791. National Bureau of Economic Research.

Ebenstein, Avraham, Lavy, Victor, and Roth, Sefi (2016). "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution". *American Economic Journal: Applied Economics* 8.4, pp. 36–65.

Fowlie, Meredith, Holland, Stephen P., and Mansur, Erin T. (2012). "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NOx Trading Program". *American Economic Review* 102.2, pp. 965–93.

Fowlie, Meredith, Rubin, Edward, and Walker, Reed (2019). "Bringing Satellite-Based Air Quality Estimates Down to Earth". *AEA Papers and Proceedings* 109, pp. 283–88.

Freeman, Richard, Liang, Wenquan, Song, Ran, and Timmins, Christopher (2019). "Willingness to pay for clean air in China". *Journal of Environmental Economics and Management* 94, pp. 188–216.

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

Graff Zivin, Joshua and Neidell, Matthew (2012). "The Impact of Pollution on Worker Productivity". *American Economic Review* 102.7, pp. 3652–73.

Grainger, C. and Ruangmas, C. (2018). "Who Wins from Emissions Trading? Evidence from California". *Environmental and Resource Economics* 71, pp. 703–727.

Harrison, Kathryn (2012). "A Tale of Two Taxes: The Fate of Environmental Tax Reform in Canada". *Review of Policy Research* 29.3, pp. 383–407.

Health Canada (2021). "Health Impacts of Air Pollution in Canada: Estimates of morbidity and premature mortality outcomes – 2021 Report".

Heckman, James J. and Smith, Jeffrey A. (1999). "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies". *The Economic Journal* 109.457, pp. 313–348.

Hernandez-Cortes, Danae and Meng, Kyle C. (2023). "Do environmental markets cause environmental injustice? Evidence from California's carbon market". *Journal of Public Economics* 217, p. 104786. ISSN: 0047-2727.

Jbaily, Abdulrahman, Zhou, Xiaodan, Liu, Jie, Lee, Ting-Hwan, Kamareddine, Leila, Verguet, Stéphane, and Dominici, Francesca (2022). "Air pollution exposure disparities across US population and income groups". *Nature* 601.7892, pp. 228–233.

Knittel, Christopher R. and Sandler, Ryan (2011). *Cleaning the Bathwater with the Baby: The Health Co-Benefits of Carbon Pricing in Transportation.* Working Papers 1115. Massachusetts Institute of Technology, Center for Energy and Environmental Policy Research.

Krewski, D. *et al.* (2009). "Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality". *Research Reports, Health Effects Institute.* Discussion 115-36.140, pp. 5–114.

Lawley, Chad and Thivierge, Vincent (2018). "Refining the evidence: British Columbia's carbon tax and household gasoline consumption". *The Energy Journal* 39.2.

Lepeule, J., Laden, F., Dockery, D., and Schwartz, J. (2012). "Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009". *Environmental Health Perspectives* 120.7, pp. 965–970.

Leroutier, Marion (2022). “Carbon pricing and power sector decarbonization: Evidence from the UK”. *Journal of Environmental Economics and Management* 111, p. 102580. ISSN: 0095-0696.

Li, Mingwei, Zhang, Da, Li, Chiao-Ting, Mulvaney, Kathleen M., Selin, Noelle E., and Karplus, Valerie J. (2018). “Air quality co-benefits of carbon pricing in China”. *Nature Climate Change* 8.5, pp. 398–403.

Liu, J. Y., Woodward, R. T., and Zhang, Y. J. (2021). “Has Carbon Emissions Trading Reduced PM2.5 in China?” *Environmental science technology* 55.10.

Menchetti, Fiammetta, Cipollini, Fabrizio, and Mealli, Fabrizia (2022). “Combining counterfactual outcomes and ARIMA models for policy evaluation”. *The Econometrics Journal* 26.1, pp. 1–24. ISSN: 1368-4221.

Meng, Jun, Li, Chi, Martin, Randall V., Donkelaar, Aaron van, Hystad, Perry, and Brauer, Michael (2019). “Estimated Long-Term (1981–2016) Concentrations of Ambient Fine Particulate Matter across North America from Chemical Transport Modeling, Satellite Remote Sensing, and Ground-Based Measurements”. en. *Environmental Science & Technology* 53.9, pp. 5071–5079.

Mildenberger, Matto, Lachapelle, Erick, Harrison, Kathryn, and Stadelmann-Steffen, Isabelle (2022). “Limited impacts of carbon tax rebate programmes on public support for carbon pricing”. *Nature Climate Change* 12.2, pp. 141–147.

Ministry of Finance (2009). *Budget and Fiscal Plan 2009/10-2011/12.* Tech. rep. British Columbia Ministry of Finance.

Ministry of Finance (2013). *Budget and Fiscal Plan 2013/14-2015/16.* Tech. rep. British Columbia Ministry of Finance.

Murray, Brian and Rivers, Nicholas (2015). “British Columbia’ s revenue-neutral carbon tax: A review of the latest ’grand experiment’ in environmental policy”. *Energy Policy* 86, pp. 674–683.

Nehiba, Cody (2022). “Correcting Heterogeneous Externalities: Evidence from Local Fuel Taxes”. en. *Journal of the Association of Environmental and Resource Economists* 9.3, pp. 495–529.

Pampalon, Robert, Hamel, Denis, Gamache, Philippe, Philibert, Mathieu D., Raymond, Guy, and Simpson, André (2012). “An Area-based Material and Social Deprivation Index for Public Health in Québec and Canada”. *Canadian Journal of Public Health / Revue Canadienne de Sante’e Publique* 103, S17–S22. (Visited on 11/09/2023).

Parry, Ian, Veung, Chandra, and Heine, Dirk (2015). “How Much Carbon Pricing is in Countries’ own Interests? The Critical Role of Co-Benefits”. *Climate Change Economics* 06.04, p. 1550019.

Persico, Claudia and Marcotte, Dave E (2022). *Air Quality and Suicide*. Working Paper 30626. National Bureau of Economic Research.

Pestel, Nico and Wozny, Florian (2021). "Health effects of Low Emission Zones: Evidence from German hospitals". *Journal of Environmental Economics and Management* 109, p. 102512.

Pretis, Felix (2022). "Does a Carbon Tax Reduce CO2 Emissions? Evidence from British Columbia". *Environmental & Resource Economics* 83.1, pp. 115–144.

Rivers, Nicholas and Schaufele, Brandon (2015). "Salience of carbon taxes in the gasoline market". *Journal of Environmental Economics and Management* 74, pp. 23–36.

Rose, Amy N., McKee, Jacob J., Sims, Kelly M., Bright, Edward A., Reith, Andrew E., and Urban, Marie L. (2020). *LandScan 2019*.

Saberian, Soodeh (2017). "The Negative Effect of the BC Carbon Tax on Vancouver Air Quality: A Good Climate for Bad Air?" *Working Paper*.

Sager, Lutz and Singer, Gregor (2022). "Clean Identification? The Effects of the Clean Air Act on Air Pollution, Exposure Disparities and House Prices". *Grantham Research Institute on Climate Change and the Environment Working Paper No. 376*, p. 57.

Sarmiento, Luis, Wägner, Nicole, and Zaklan, Aleksandar (2022). "The Air Quality and Well-Being Effects of Low Emission Zones". en. *SSRN Electronic Journal*.

Shapiro, Joseph S and Walker, Reed (2021). *Where is Pollution Moving? Environmental Markets and Environmental Justice*. Working Paper 28389. National Bureau of Economic Research.

Sheriff, Glenn (2023). "California's GHG cap and trade program and the equity of air toxic releases". *Journal of the Association of Environmental and Resource Economists* Forthcoming.

Shindell, Drew T., Lee, Yunha, and Faluvegi, Greg (2016). "Climate and health impacts of US emissions reductions consistent with 2°C". *Nature Climate Change* 6.5, pp. 503–507.

Statistics Canada (2021a). *Mortality and potential years of life lost, by selected causes of death and sex, three-year average, census metropolitan areas. Table 13-10-0741-01.*

Statistics Canada (2021b). *Supply and disposition of refined petroleum products, monthly. Table 25-10-0044-01.*

Statistics Canada (2021c). *Upper income limit, income share and average of adjusted market, total and after-tax income by income decile. Table 11-10-0193-01.*

Vandyck, Toon, Keramidas, Kimon, Kitous, Alban, Spadaro, Joseph, Van Dingenen, Rita, Holland, Mike, and Saveyn, Bert (2018). "Air quality co-benefits for human health and agriculture counterbalance costs to meet Paris Agreement pledges". *Nature Communications* 9.

Wagner, Ulrich and Preux, Laure de (2016). *The co-benefits of climate policy: evidence from the EU Emissions Trading Scheme*. VfS Annual Conference 2016 (Augsburg): Demographic Change. Verein für Socialpolitik / German Economic Association.

Wen, Jeff and Burke, Marshall (2022). "Wildfire Smoke Exposure Worsens Learning Outcomes". *Nature Sustainability* 5, pp. 920–921.

West, Jason, Smith, Steven, Silva, Raquel, Naik, Vaishali, Zhang, Yuqiang, Adelman, Zachariah, Fry, Meridith, Anenberg, Susan, Horowitz, Larry, and Lamarque, Jean-François (2013). "Co-benefits of Global Greenhouse Gas Mitigation for Future Air Quality and Human Health". *Nature Climate Change* 3.

Zhang, Wen-Wen, Zhao, Bin, Ding, Dian, Sharp, Basil, Gu, Yu, Xu, Shi-Chun, Xing, Jia, Wang, Shu-Xiao, Liou, Kuo-Nan, and Rao, Lan-Lan (2021). "Co-benefits of subnationally differentiated carbon pricing policies in China: Alleviation of heavy PM2.5 pollution and improvement in environmental equity". *Energy Policy* 149, p. 112060.

Zhu, Junming, Li, Xiao, Fan, Yichun, Shi, Haijia, and Zhao, Lei (2022). "Effect of carbon market on air pollution: Firm-level evidence in China". *Resources, Conservation and Recycling* 182, p. 106321. ISSN: 0921-3449.

van Donkelaar, Aaron, Martin, Randall V., Li, Chi, and Burnett, Richard T. (2019). "Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors". *Environmental Science & Technology* 53.5, pp. 2595–2611.

von Bergmann, Jens (2021). *tongfen: R package to Make Data Based on Different Geographies Comparable*. R package version 0.3.3.

von Bergmann, Jens, Shkolnik, Dmitry, and Jacobs, Aaron (2022). *cancensus: R package to access, retrieve, and work with Canadian Census data and geography*. R package version 0.5.3.

A. Descriptive Statistics

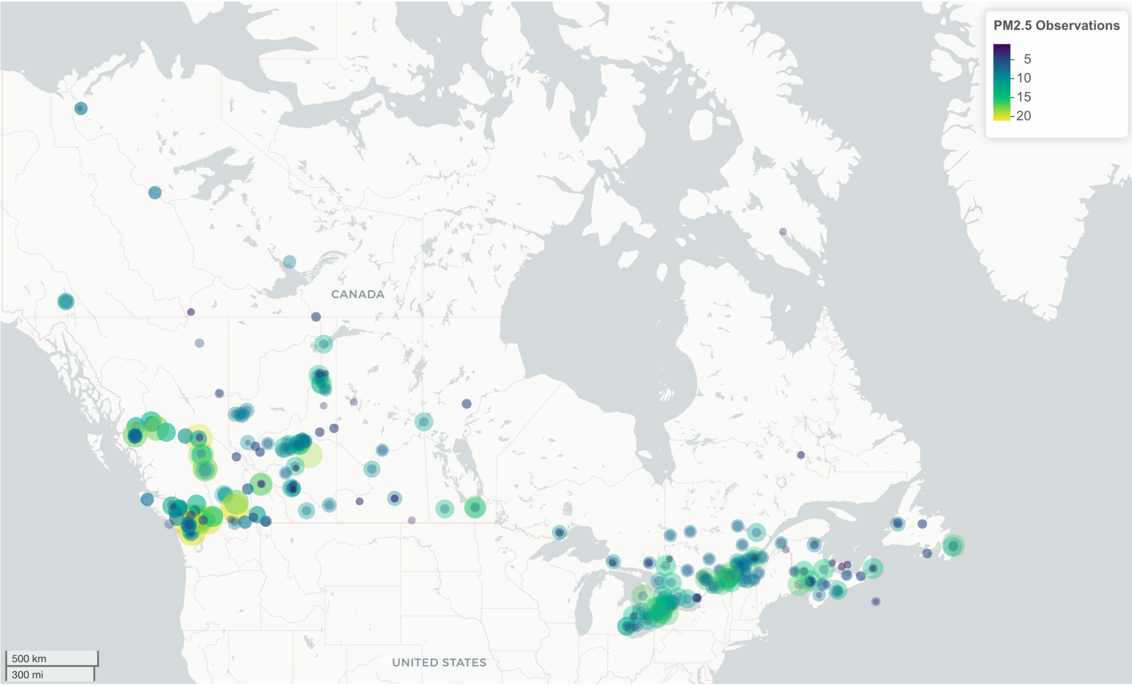

Figure A.1: Availability of $PM_{2.5}$ readings in the National Atmospheric Surveillance Program database between 2000 and 2018. Lighter colours and larger dot sizes indicate higher availability of readings (monitoring stations which were added earlier).

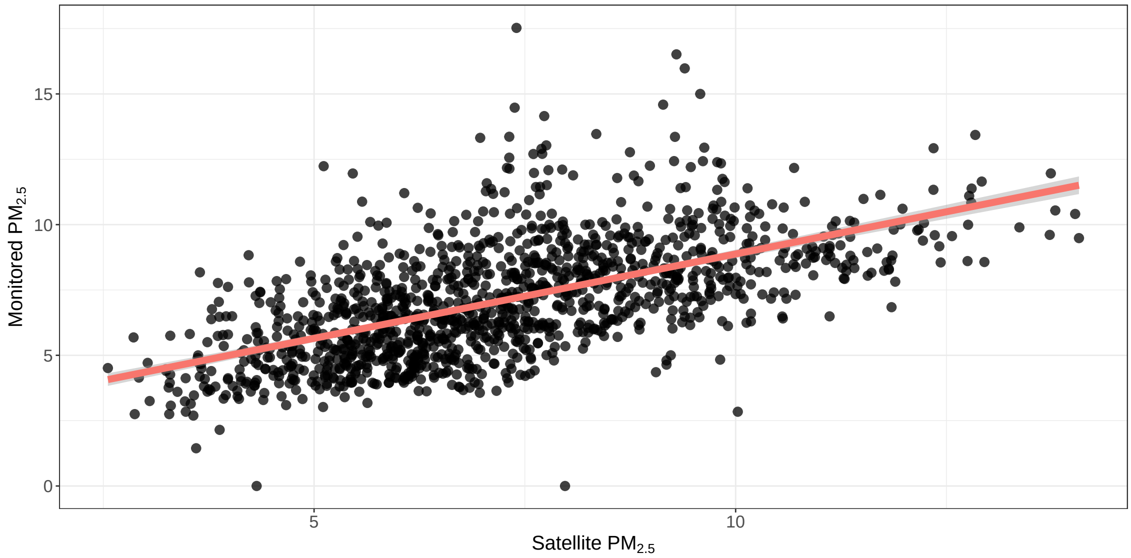

Figure A.2: Scatterplot of satellite $PM_{2.5}$ (Meng *et al.*, 2019) (y-axis) and $PM_{2.5}$ from NAPS monitoring stations (x-axis). Both measures are in $\mu g/m^3$. The correlation coefficient is 0.597.

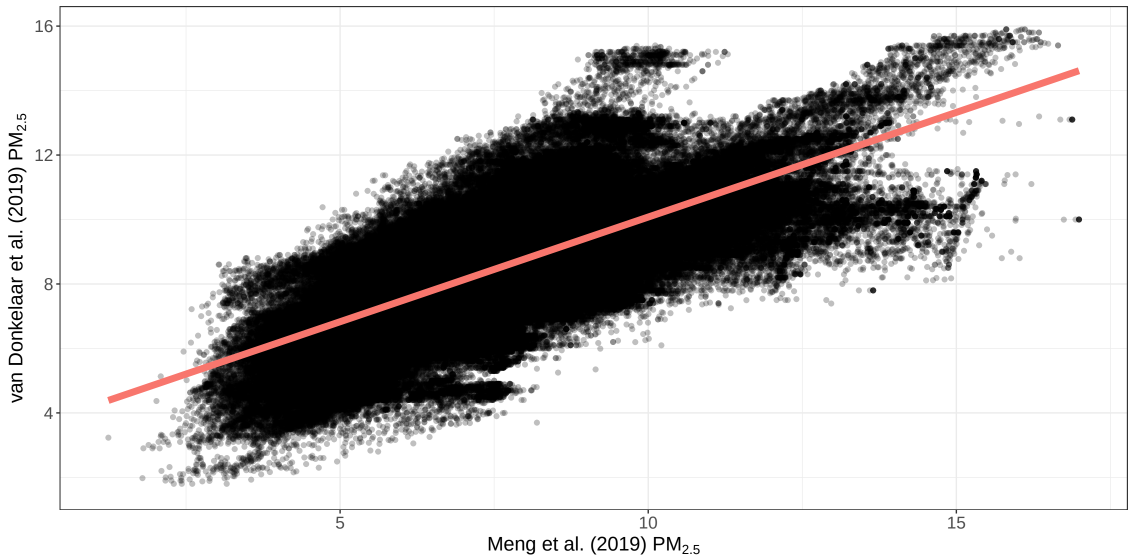

Figure A.3: Scatterplot of satellite $PM_{2.5}$ (Meng *et al.*, 2019) (x-axis) vs Satellite $PM_{2.5}$ (van Donkelaar *et al.*, 2019) (y-axis). Both measures are in $\mu g/m^3$. The correlation coefficient is 0.729.

Table A.1: Summary Statistics, 2000-2007

	Control Provinces			British Columbia		
Variable	N	Mean	SD	N	Mean	SD
$PM_{2.5}$ (van Donkelaar *et al.*, 2019)	175870	9.52	1.54	27920	8.06	1.19
$PM_{2.5}$ (Meng *et al.*, 2019)	175870	8.61	2.07	27920	6.95	1.39
Pop. Density (Rose *et al.*, 2020)	175912	3358.26	3375.33	27920	3169.94	2136.98
Median Income	43978	26341.65	9088.59	6980	25055.65	8090.75
Material Deprivation Index	20606	46.48	28.60	3313	43.57	28.12
High Emission Commute %	43701	74.93	18.33	6926	77.64	16.37
Low Emission Commute %	43701	24.52	18.26	6926	21.62	16.26
Public Transport Commute %	43701	17.02	14.48	6926	13.06	10.68
Zero Emission Commute %	43701	7.50	9.43	6926	8.56	10.79
Precipitation (Abatzoglou *et al.*, 2018)	175768	74.05	21.74	27920	131.20	37.06
Max Temperature (Abatzoglou *et al.*, 2018)	175768	11.93	1.58	27920	14.55	0.66
Min Temperature (Abatzoglou *et al.*, 2018)	175768	1.74	2.50	27920	6.46	0.62
Wind Speed (Abatzoglou *et al.*, 2018)	175768	3.63	0.49	27920	2.98	0.16

Table A.2: Summary Statistics, 2008-2018

	Control Provinces			British Columbia		
Variable	N	Mean	SD	N	Mean	SD
$PM_{2.5}$ (van Donkelaar *et al.*, 2019)	241865	8.15	1.51	38390	6.09	0.95
$PM_{2.5}$ (Meng *et al.*, 2019)	197888	7.35	1.73	31410	6.07	1.10
Population Density (Rose *et al.*, 2020)	241879	3614.39	3365.36	38390	3478.58	2305.69
Median Income	43978	33324.06	11718.48	6980	31772.63	9765.79
High Emission Commute %	43806	74.39	20.20	6955	72.94	18.75
Low Emission Commute %	43806	25.06	20.12	6955	26.25	18.61
Public Transport Commute %	43806	18.85	15.89	6955	18.38	13.38
Zero Emission Commute %	43806	6.21	10.15	6955	7.87	11.32
Precipitation (Abatzoglou *et al.*, 2018)	241681	77.57	21.99	38390	134.58	37.33
Max Temperature (Abatzoglou *et al.*, 2018)	241681	12.32	1.76	38390	14.58	0.86
Min Temperature (Abatzoglou *et al.*, 2018)	241681	2.11	2.59	38390	6.56	0.80
Wind Speed (Abatzoglou *et al.*, 2018)	241681	3.64	0.48	38390	3.00	0.19

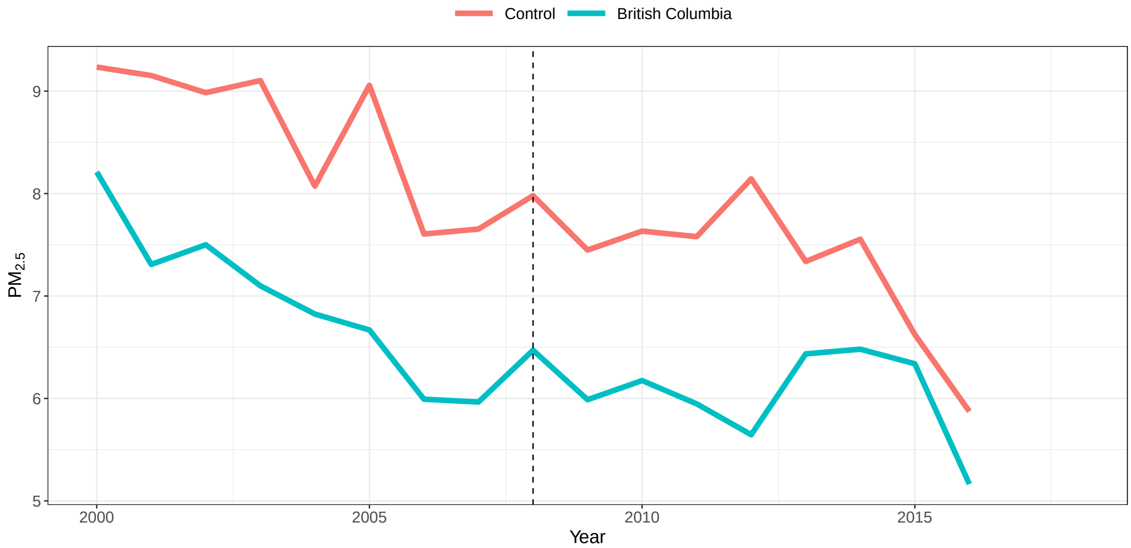

Figure A.4: Trends in satellite $PM_{2.5}$ (Meng *et al.*, 2019), British Columbia and average of control provinces, between 2000 and 2016. The implementation of the carbon tax in 2008 is highlighted by the dashed vertical line.

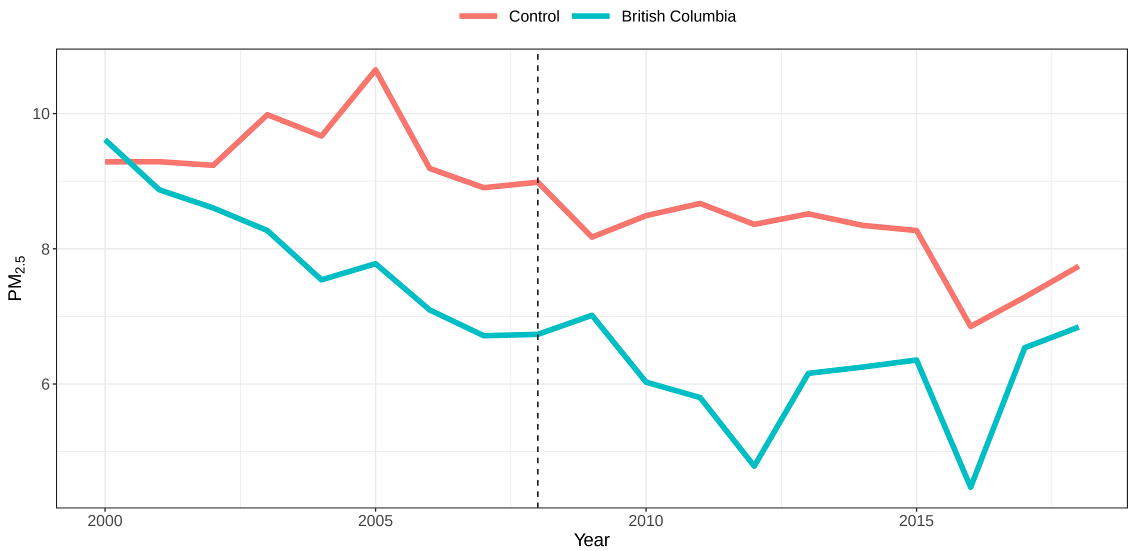

Figure A.5: Trends in satellite $PM_{2.5}$ (van Donkelaar *et al.*, 2019), British Columbia and average of control provinces, between 2000 and 2018. The implementation of the carbon tax in 2008 is highlighted by the dashed vertical line.

Figure A.6: Event study aggregation of a TWFE-DID regression with $PM_{2.5}$ (Meng *et al.*, 2019) as the dependent variable. Event study coefficients reported in black; aggregate pre- and post-treatment effects reported in red.

B. Comparison between TWFE-DID, SCM and SDID

In order to formally explain how SDID combines features from TWFE-DID and SCM, let me consider a balanced panel with N observations and T time periods. In the British Columbian case, the outcome variable is $PM_{2.5it}$, and the binary treatment is TAX_{it}. Let $i = 1, ..., N_{tr}$ be the treated DAs in BC, and $i = N_{tr} + 1, ..., N_{co}$ be the DAs in control provinces. The baseline TWFE-DID regression problem can be expressed as:

$$(\hat{\tau}^{did}\hat{\mu}, \hat{\eta}, \hat{\theta}) = \underset{\tau,\mu,\eta,\theta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (PM_{2.5it} - \mu - \eta_i - \theta_t - \tau TAX_{it})^2 \right\} \tag{9}$$

Which is solved without the use of unit or time-specific weights, but with the inclusion of unit and time-specific fixed effects η_i and θ_t as also illustrated in Equation 2. The SCM estimator, instead, does not employ unit fixed effects, but includes time fixed effects and unit-specific weights ω_i^{sc}:

$$(\hat{\tau}^{sc}, \hat{\mu}, \hat{\eta}, \hat{\theta}) = \underset{\tau,\mu,\eta,\theta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (PM_{2.5it} - \mu - \theta_t - \tau TAX_{it}) \hat{\omega}_i^{sc} \right\} \tag{10}$$

Finally, the SDID estimator combines features from Equation 9 and Equation 10. Unit weights $\hat{\omega}_i^{sdid}$ are chosen such that the pre-treatment outcome path of control DAs are parallel to those of the treated units[40]:

$$\omega_0 + \sum_{i=N_{tr}+1}^{N_{co}} \hat{\omega}_i^{sdid} PM_{2.5it} \approx \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} PM_{2.5it} \tag{11}$$

Moreover, time weights $\hat{\lambda}_t^{sdid}$ need to ensure that the pre-treatment levels for the control units differs from the post-treatment levels for the same units only by a constant. Letting $t = 1, ..., T$ be the total length of the panel, T_{pre} be the number of pre-intervention periods, and T_{post} be the number of post-intervention periods, the condition can be expressed as:

[40]Unit-specific weights are found using a regularisation parameter ζ, as in Doudchenko and Imbens (2016), which aids the estimation strategy by increasing the dispersion of the weights and ensuring their uniqueness. When the intercept ω_0 and the regularisation parameter are set to 0, the unit weights ω_i correspond to the SCM weights in Abadie *et al.* (2010). For further details on the procedure used to estimate ζ, please refer to Arkhangelsky *et al.* (2021).

$$\lambda_0 + \sum_{t=1}^{T_{pre}} \hat{\lambda}_t^{sdid} PM_{2.5it} \approx \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} PM_{2.5it} \tag{12}$$

Thus, the regression problem for the SDID estimator can be expressed as a weighted TWFE-DID problem which incorporates unit and time-specific fixed effects η_i and θ_t, plus unit and time-specific weights ω_i and λ_t, as illustrated in Equation 13:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\eta}, \hat{\theta}) = \underset{\tau,\mu,\eta,\theta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \eta_i - \theta_t - \tau TAX_{it})^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \tag{13}$$

C. Additional Details on Mechanisms

C.1. Fuel Demand

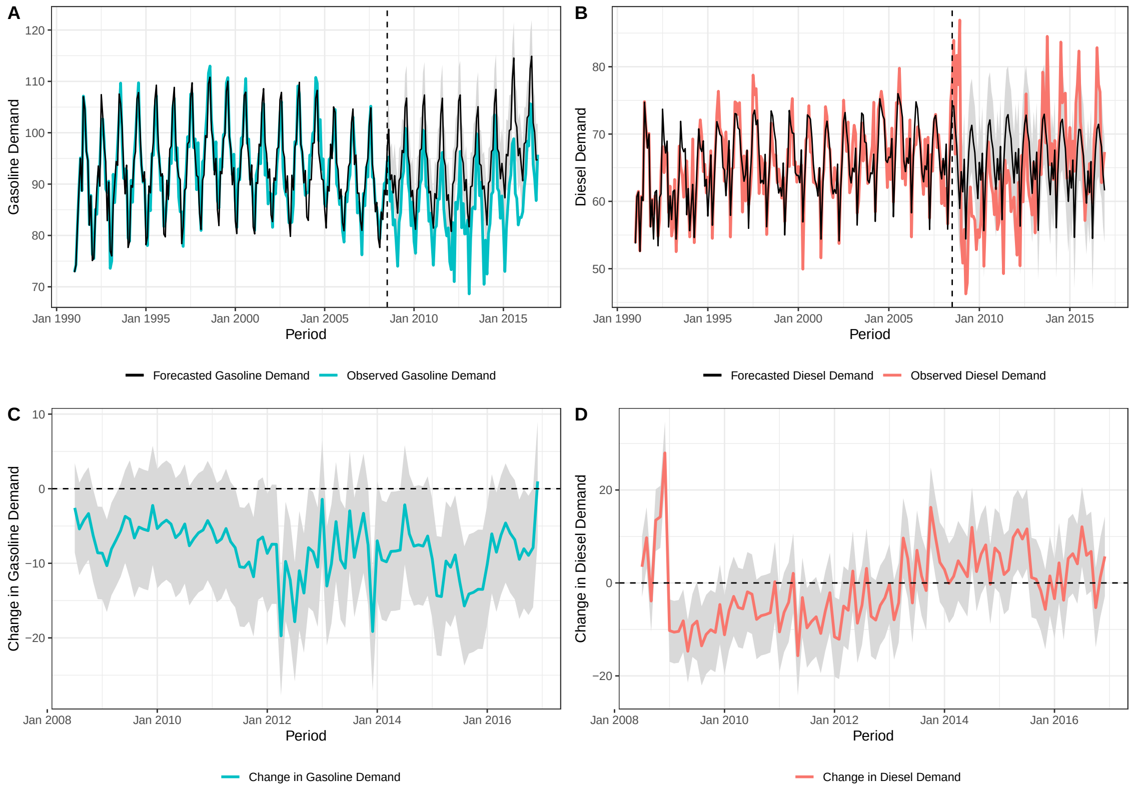

Figure C.1: Graphical results from the C-ARIMA regressions on monthly gasoline and diesel sales. Panel (A) and (B) show the observed and forecasted gasoline and diesel sales time series for the full post-intervention horizon. Panels (C) and (D) represent the gap between observed and forecasted series for gasoline and diesel sales.

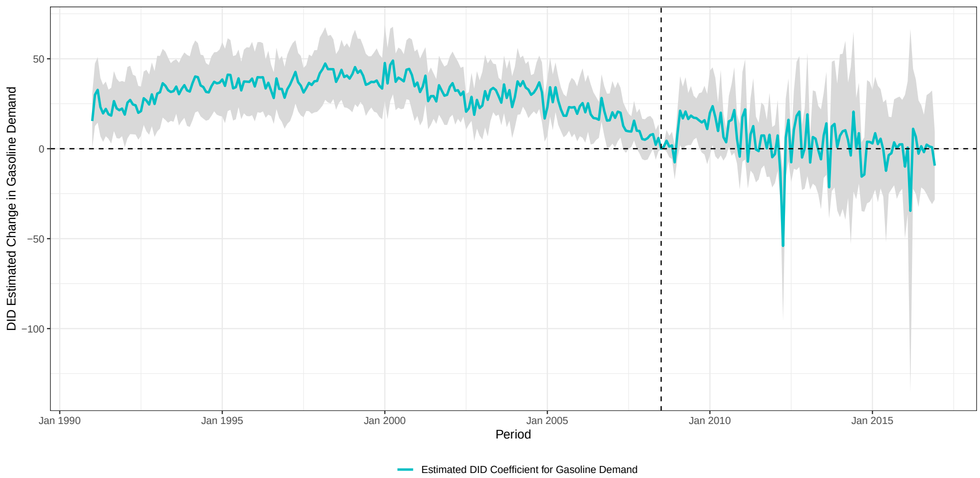

Figure C.2: Event study regression of monthly gasoline sales at the province level.

C.2. Commute Mode Switching Empirical Analysis

Ideally, when concerned with the estimation of $PM_{2.5}$ reductions arising from the implementation of carbon pricing, I would look at DA-level reductions in motor fuel sales or in the quantity of vehicle kilometres travelled; however, these data are not available at the desired level of granularity for Canada between 2000 and 2018. The only precedent of a paper studying the relationship between the 2008 BC carbon tax and air quality in Canadian cities is the working paper of Saberian (2017), who restricts the analysis to Vancouver and uses monitoring stations data in order to infer her result – pointing to a worsening in air pollution following the carbon tax. The analysis of mechanisms leading to the result in Saberian (2017) highlights gasoline-to-diesel fuel switching as the potential causal driver of increased air pollution. However, the evidence is only anecdotal, as no evidence supporting the claim is presented in the study. Moreover, while Canadian province-level data on vehicle sales disaggregated by type of fuel is only available from 2011 onwards, the post-2011 trends in sales of diesel vehicles are relatively flat (See Figure C.4), and the landscape seems to be dominated by gasoline cars (See Figure C.3), suggesting that an eventual gas-to-diesel switch caused by the carbon tax incentive would have produced all of its results between July 2008 and January 2011 before bottoming out; the evidence for this conclusion is not very strong as a result. Another potential mechanism behind an increase in air pollution could derive from an exceptionally high rate of replacement in BC's car fleet with respect to other Canadian provinces, caused by the willingness of BC's residents to increase their cars' fuel efficiency and realise savings at the pump. If the savings per each tank refuel were sufficient to offset the increase in gasoline prices due to the carbon tax, British Columbian residents could have potentially travelled more kilometres than prior to the tax, thereby increasing road congestion and hence pollution due to a rebound effect. As shown in Figure C.5 there has indeed been a rapid increase in truck and SUV sales in British Columbia after 2008; however, this increase is paralleled by similar jumps in truck sales in all large Canadian provinces[41], and it thus seems implausible to attribute it to the marginal effect of the carbon tax in raising fuel prices.

I instead exploit the information contained in the 2001, 2006, 2011 and 2016 waves of the Canadian census, which contains data on commute-to-work modes at the DA level for all Canadian CMAs. While the information on commute modes is not an exhaustive representation of all car trips made in each DA, the granularity of the data may shed light on whether residents of DAs located in British Columbia have

[41] Namely, Alberta, Ontario and Quebec.

adjusted their behaviour following the implementation of the carbon tax, substituting public transport or active commuting modes such as cycling and walking for car trips. In particular, I estimate the following equation:

$$Mode_{it} = \tau TAX_{it} + \theta_t + \eta_i + \epsilon_{it} \tag{14}$$

Where $Mode_{it}$ is the share of each commute mode (high emission, low emission, public transport and zero emission), Tax_{it} is the carbon tax DID binary variable, θ_t and η_i are time and unit-specific fixed effects, and ϵ_{it} is an idiosyncratic error term. In additional specifications, I also add a vector of controls X_{it} which account for population density, median income, and weather covariates (precipitation, maximum and minimum temperature, and wind speed), hence the estimating equation becomes:

$$Mode_{it} = \tau TAX_{it} + \beta X_it + \theta_t + \eta_i + \epsilon_{it} \tag{15}$$

I initially run the TWFE-DID regressions for the whole sample, without trimming the control pool. In further specifications, I restrict the control sample to the units which receive positive ω_i weights in the SDID estimation of the main result, in order to ensure comparability across treatment and control cohorts and reduce the reliance on potentially violated parallel trends. Further, I retrieve the ω_i weights from the SDID estimation and weigh my restricted TWFE-DID regressions with the SDID weights, assigning equal weights $\frac{1}{N_{tr}}$ to the treatment cohort.

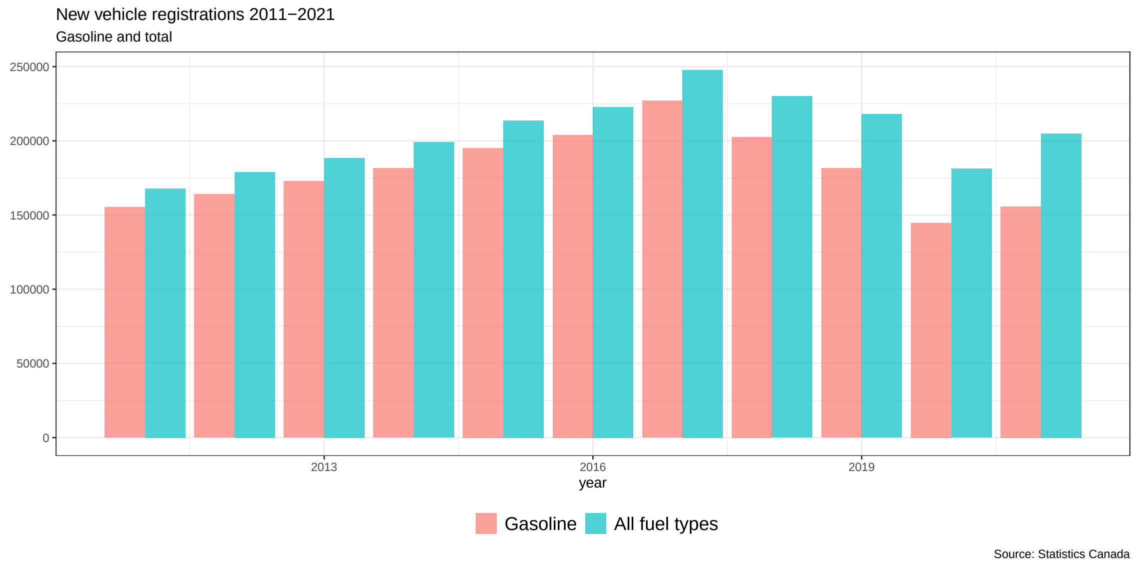

Figure C.3: New vehicle registrations in BC, 2011-2021: gasoline and all other fuel types.

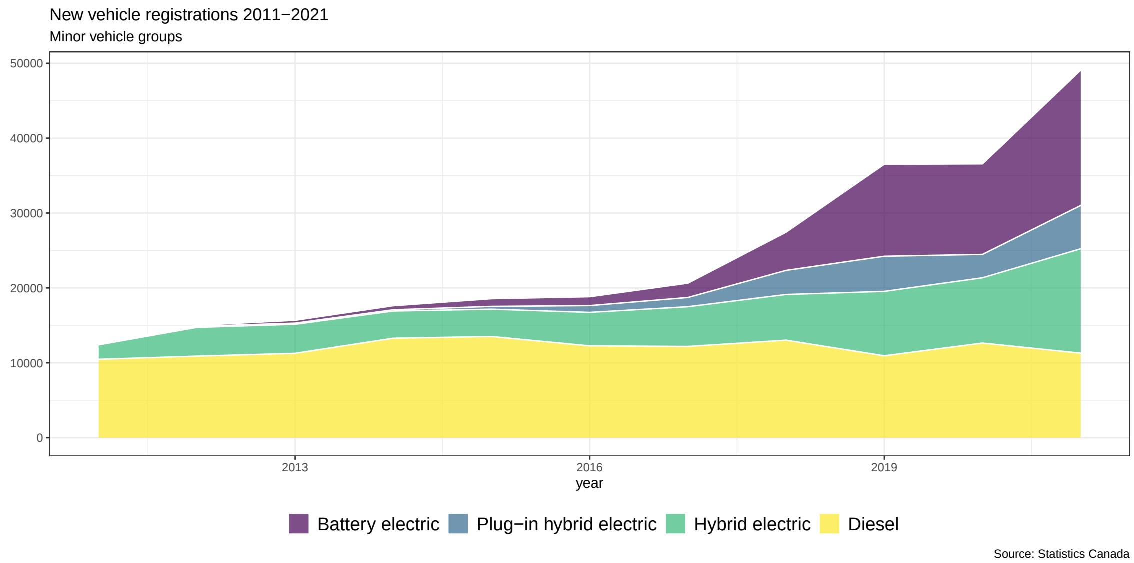

Figure C.4: New vehicle registrations in BC, 2011-2021: all other fuel types.

Figure C.5: Passenger cars vs Truck and SUV sales, large Canadian Provinces, 1990-2021.

Table C.1: TWFE-DID results for low emissions commute mode

	Low Emission Commute Mode					
	(1)	(2)	(3)	(4)	(5)	(6)
DID	0.0408***	0.0516***	0.0535***	0.0457***	0.0510***	0.0506***
	(0.0111)	(0.0103)	(0.0110)	(0.0109)	(0.0113)	(0.0114)
DA FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls				✓	✓	✓
SDID control pool		✓	✓		✓	✓
SDID weights			✓			✓
R^2	0.87321	0.84174	0.84532	0.87715	0.84674	0.84996
Adjusted R^2	0.83078	0.78876	0.79354	0.83560	0.79490	0.79920
Observations	101,358	38,769	38,769	100,244	38,348	38,348

Notes: The dependent variable is the dissemination area level share of low emissions commutes. All regressions include dissemination area and year fixed effects. Columns (4)-(6) include controls for precipitation, maximum and minimum temperature, and wind speed, plus the natural logarithm of population and median income. Columns (2), (3), (5) and (6) restrict the control unit pool to DAs which receive positive weights in the main SDID regression. Columns (3) and (6) additionally include the estimated SDID unit weights ω_i as regression weights. Standard errors are clustered at the CMA level. ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$

Table C.2: TWFE-DID results for zero emissions commute mode

	Zero Emission Commute Mode					
	(1)	(2)	(3)	(4)	(5)	(6)
DID	0.0057**	0.0106***	0.0117***	0.0066***	0.0088***	0.0092***
	(0.0025)	(0.0017)	(0.0021)	(0.0022)	(0.0016)	(0.0016)
DA FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Controls				✓	✓	✓
SDID control pool		✓	✓		✓	✓
SDID weights			✓			✓
R^2	0.80811	0.80808	0.81877	0.81200	0.81355	0.82463
Adjusted R^2	0.74390	0.74383	0.75810	0.74841	0.75047	0.76531
Observations	101,358	38,769	38,769	100,244	38,348	38,348

Notes: The dependent variable is the dissemination area level share of zero emissions commutes. All regressions include dissemination area and year fixed effects. Columns (4)-(6) include controls for precipitation, maximum and minimum temperature, and wind speed, plus the natural logarithm of population and median income. Columns (2), (3), (5) and (6) restrict the control unit pool to DAs which receive positive weights in the main SDID regression. Columns (3) and (6) additionally include the estimated SDID unit weights ω_i as regression weights. Standard errors are clustered at the CMA level. ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$

D. Robustness Checks

D.1. Composition of SDID Control Pools

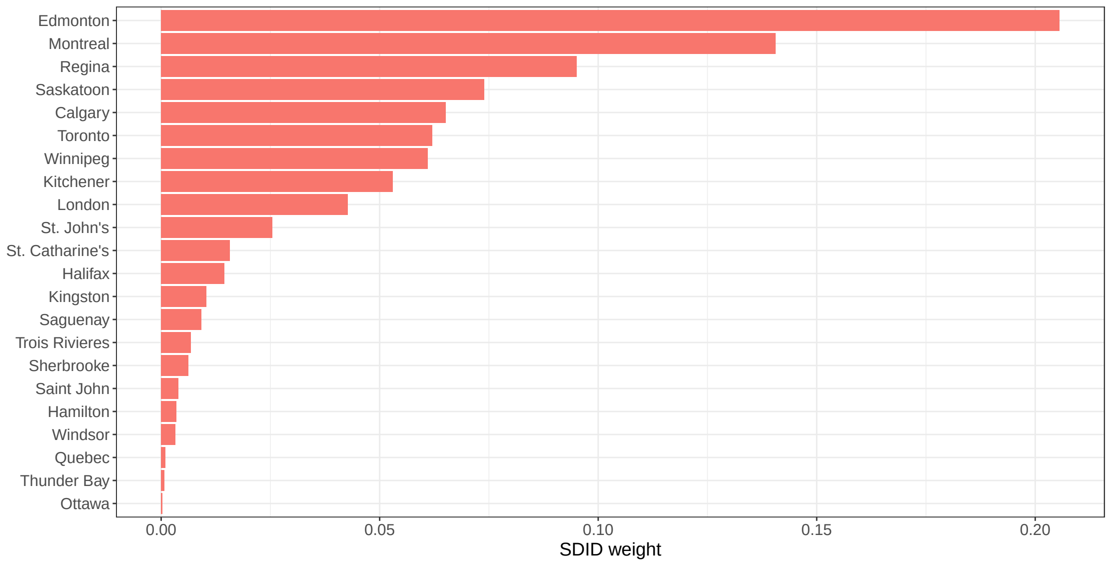

Figure D.1: Composition of the synthetic DID unit of Figure 2. Individual DA weights are aggregated up to the CMA level.

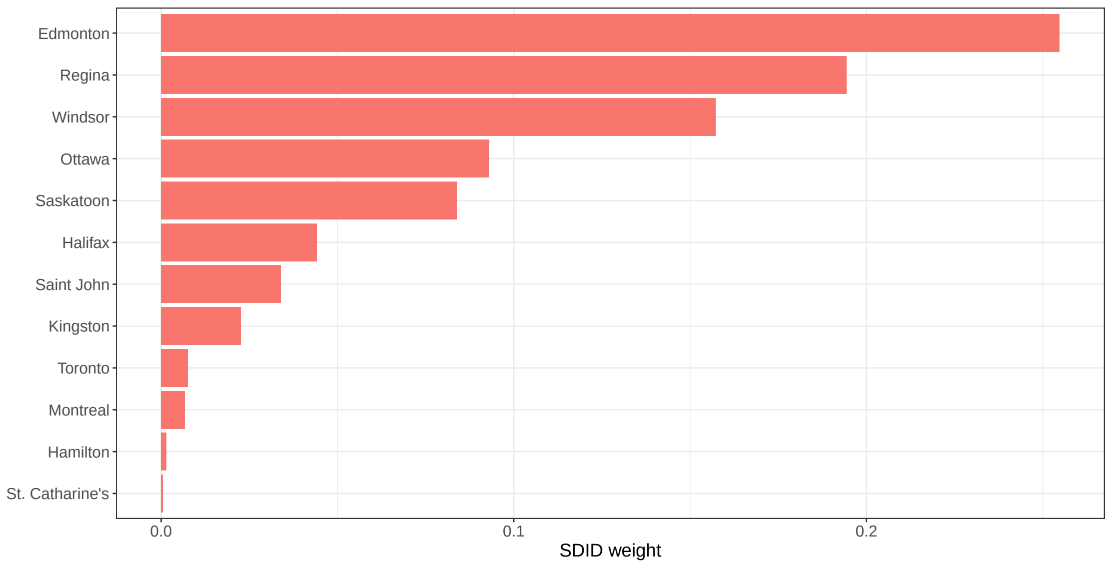

Figure D.2: Composition of the synthetic DID unit of Figure D.3. Individual DA weights are aggregated up to the CMA level.

D.2. Main results with van Donkelaar *et al.* (2019) $PM_{2.5}$ data

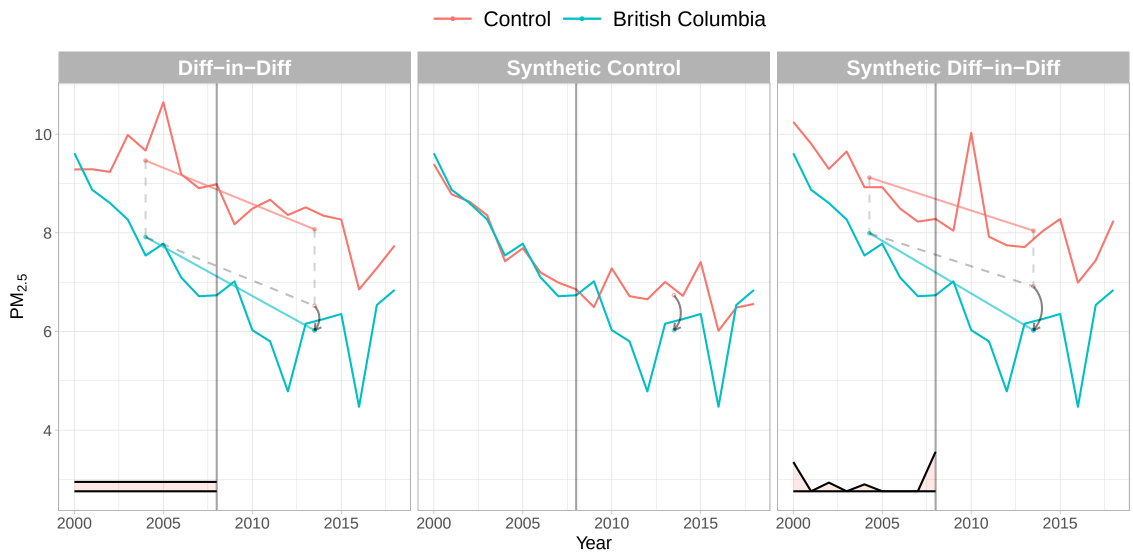

Figure D.3: Graphical results from DID, SCM and SDID for $PM_{2.5}$ concentrations, with van Donkelaar *et al.* (2019) data. Time weights λ_t are represented in light red at the bottom of the pre-intervention panel. The curved arrows graphically represent the ATT over the post-intervention period.

Table D.1: Summary of $\hat{\tau}$ point estimates and standard errors from all estimation methods, dependent variable from van Donkelaar *et al.* (2019).

	(1) ***DID***	(2) ***SCM***	(3) ***SDID***
$\hat{\boldsymbol{\tau}}$	-0.4954	-0.7087	-0.8896
	(0.0085)	(0.1540)	(0.0300)
Unit FE	✓		✓
Year FE	✓	✓	✓
ω_i		✓	✓
λ_t			✓
$\boldsymbol{N_{obs}}$	483873	483873	483873

Notes: All point estimates represent the average impact of the 2008 carbon tax during the 2009-2016 post-treatment period. Standard errors are calculated using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021) with 200 replications. All regressions use 2000-2016 data.

D.3. DAs in the Vancouver CMA

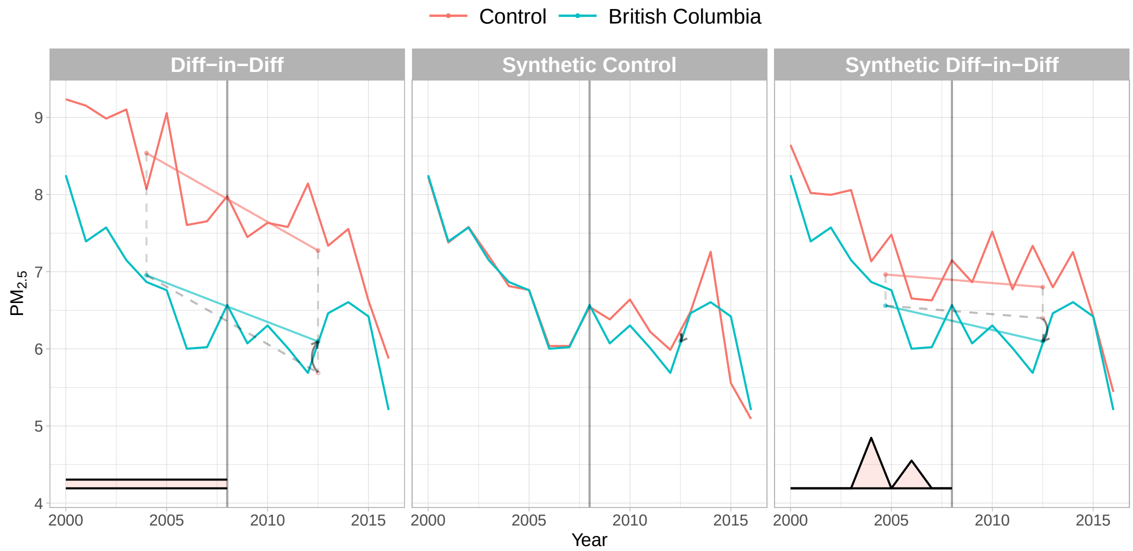

Figure D.4: Graphical results from DID, SCM and SDID for $PM_{2.5}$ concentrations, with Meng *et al.* (2019) data, dataset restricted to DAs in the Vancouver CMA.

Table D.2: Summary of $\hat{\tau}$ point estimates and standard errors from all estimation methods, dependent variable from Meng *et al.* (2019), dataset restricted to DAs in the Vancouver CMA.

	(1) ***DID***	(2) ***SCM***	(3) ***SDID***
$\hat{\boldsymbol{\tau}}$	0.4061	-0.1062	-0.3014
	(0.0071)	(0.0702)	(0.0225)
Unit FE	✓		✓
Year FE	✓	✓	✓
ω_i		✓	✓
λ_t			✓
$\boldsymbol{N_{obs}}$	422467	422467	422467

Notes: All point estimates represent the average impact of the 2008 carbon tax during the 2009-2016 post-treatment period. Standard errors are calculated using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021) with 200 replications. All regressions use 2000-2016 data.

D.4. DAs matching NAPS Monitoring Stations

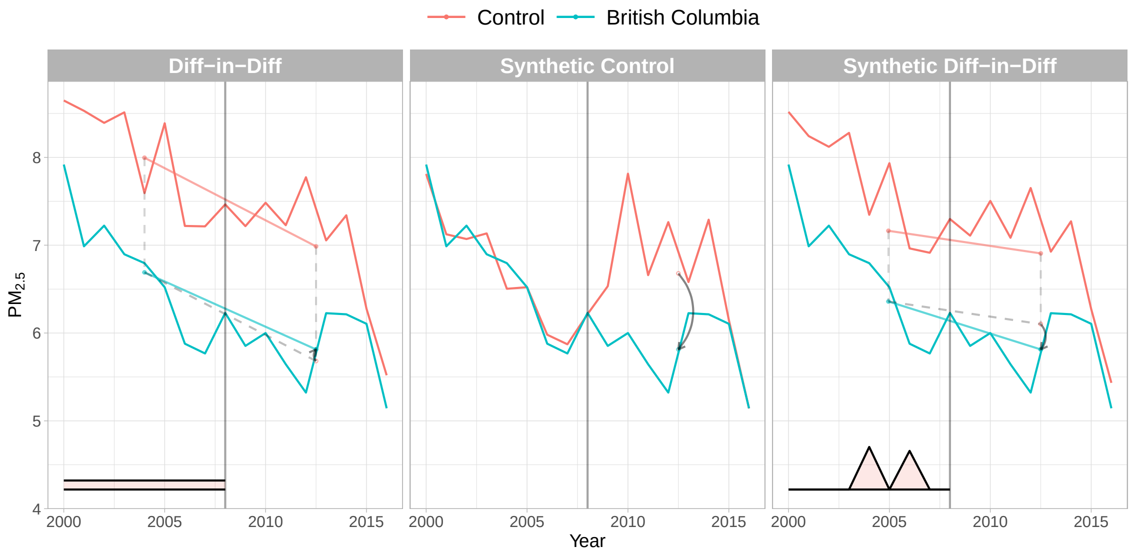

Figure D.5: Graphical results from DID, SCM and SDID for $PM_{2.5}$ concentrations, with Meng *et al.* (2019) data, dataset restricted to DAs matching NAPS monitoring stations' locations.

Table D.3: Summary of $\hat{\tau}$ point estimates and standard errors from all estimation methods, dependent variable from Meng *et al.* (2019), dataset restricted to DAs matching NAPS monitoring stations' locations.

	(1) ***DID***	(2) ***SCM***	(3) ***SDID***
$\hat{\boldsymbol{\tau}}$	0.132	-0.865	-0.288
	(0.117)	(0.128)	(0.097)
Unit FE	✓		✓
Year FE	✓	✓	✓
ω_i		✓	✓
λ_t			✓
$\boldsymbol{N_{obs}}$	2227	2227	2227

Notes: All point estimates represent the average impact of the 2008 carbon tax during the 2009-2016 post-treatment period. Standard errors are calculated using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021) with 200 replications. All regressions use 2000-2016 data.

D.5. Post-treatment period limited to 2013

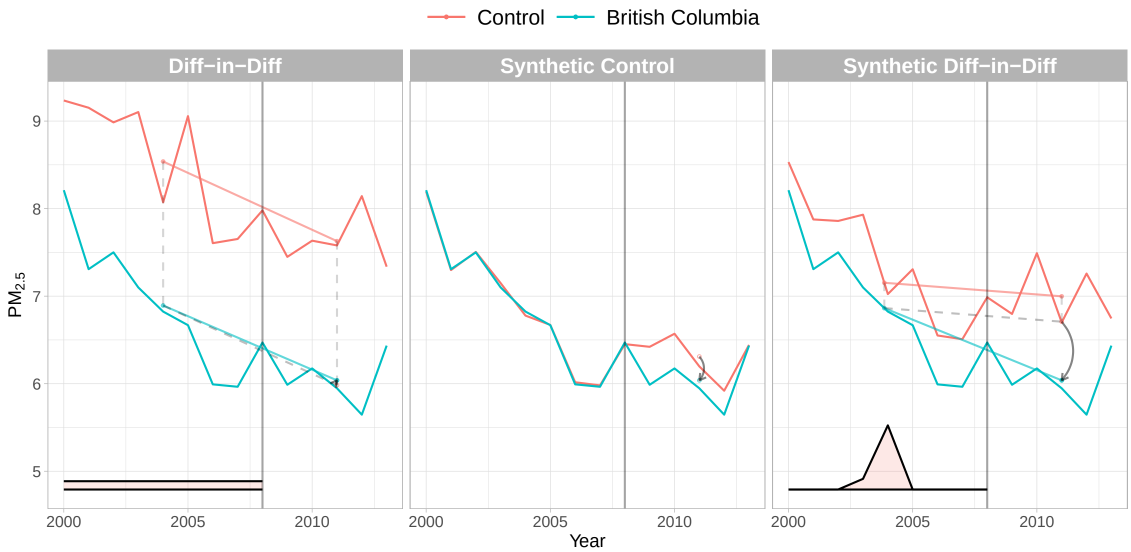

Figure D.6: Graphical results from DID, SCM and SDID for $PM_{2.5}$ concentrations, with Meng *et al.* (2019) data, dataset restricted to 2013.

Table D.4: Summary of $\hat{\tau}$ point estimates and standard errors from all estimation methods, dependent variable from Meng *et al.* (2019), dataset restricted to 2013.

	(1) ***DID***	(2) ***SCM***	(3) ***SDID***
$\hat{\boldsymbol{\tau}}$	0.0547 (0.0081)	-0.2723 (0.0803)	-0.6703 (0.0341)
Unit FE	✓		✓
Year FE	✓	✓	✓
ω_i		✓	✓
λ_t			✓
$\boldsymbol{N_{obs}}$	432939	432939	432939

Notes: All point estimates represent the average impact of the 2008 carbon tax during the 2009-2013 post-treatment period. Standard errors are calculated using the bootstrap variance estimation algorithm described in Arkhangelsky *et al.* (2021) with 200 replications. All regressions use 2000-2013 data.

D.6. TWFE-DID using NAPS Monitoring Stations Data

Table D.5: DID with NAPS monitors, 1998-2016

	η $PM_{2.5}$ (1)	μ $PM_{2.5}$ (2)	η $PM_{2.5}$ (3)	μ $PM_{2.5}$ (4)
Carbon Tax	0.0396	0.0364	0.0248	0.0001
	(0.0314)	(0.0219)	(0.0180)	(0.0122)
City FE	✓	✓	✓	✓
Month*Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Sample	Top 15	Top 15	Full	Full
Observations	3,372	3,329	16,116	15,249
Adjusted R^2	0.44870	0.57613	0.32740	0.51854
Dependent variable mean	4.0168	7.1371	3.9793	7.1407

Notes: η is the number of $PM_{2.5}$ monthly safe threshold exceedances. μ is the average monthly $PM_{2.5}$ concentration. Controls include the retail price of gasoline and diesel, unemployment rate, after tax income, maximum and minimum temperature, precipitation and wind speed. Top 15 cities are Vancouver, Calgary, Edmonton, Halifax, Hamilton, Ottawa, Saskatoon, St. Johns, Toronto, Windsor, Winnipeg, Kitchener, London, St. Catharines, Oshawa and Regina. Full panel includes all NAPS air quality monitoring stations. Standard errors clustered at the City level. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table D.6: DID with NAPS monitors, 1998-2013

	η $PM_{2.5}$ (1)	μ $PM_{2.5}$ (2)	η $PM_{2.5}$ (3)	μ $PM_{2.5}$ (4)
Carbon Tax	0.0646**	0.0422**	0.0196	0.0030
	(0.0299)	(0.0185)	(0.0240)	(0.0160)
City FE	✓	✓	✓	✓
Month*Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Sample	Top 15	Top 15	Full	Full
Observations	2,796	2,764	12,744	12,006
Adjusted R^2	0.49176	0.61504	0.33748	0.54680
Dependent variable mean	4.0303	7.0618	4.1321	7.1454

Notes: η is the number of $PM_{2.5}$ monthly safe threshold exceedances. μ is the average monthly $PM_{2.5}$ concentration. Controls include the retail price of gasoline and diesel, unemployment rate, after tax income, maximum and minimum temperature, precipitation and wind speed. Top 15 cities are Vancouver, Calgary, Edmonton, Halifax, Hamilton, Ottawa, Saskatoon, St. Johns, Toronto, Windsor, Winnipeg, Kitchener, London, St. Catharines, Oshawa and Regina. Full panel includes all NAPS air quality monitoring stations. Standard errors clustered at the City level. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

E. Pollution, Health and Distributional Implications

E.1. Estimates using RR from Krewski *et al.* (2009)

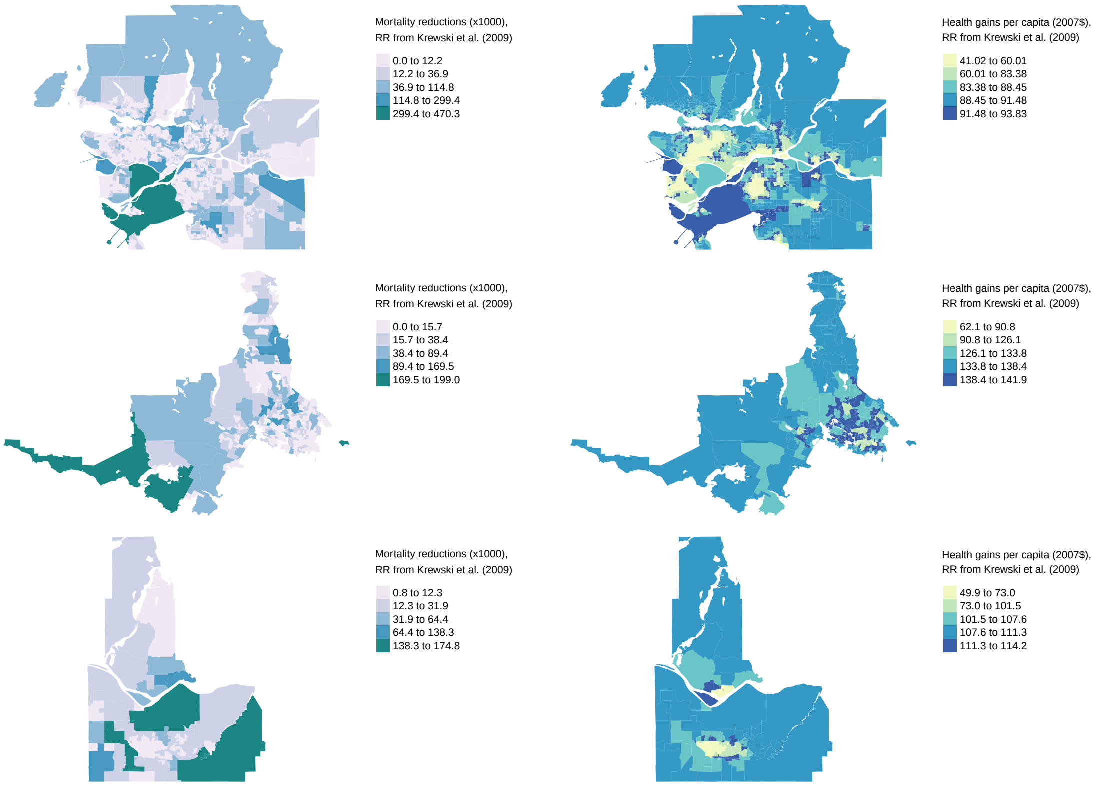

Figure F.1: Spatial distribution of mortality reductions per 1000 residents (left panel) and health gains per capita (right panel) using the RR estimates from Krewski *et al.* (2009), for the Vancouver (top row), Victoria (middle row) and Abbotsford (bottom row) CMAs.

E.2. Health-income relationships

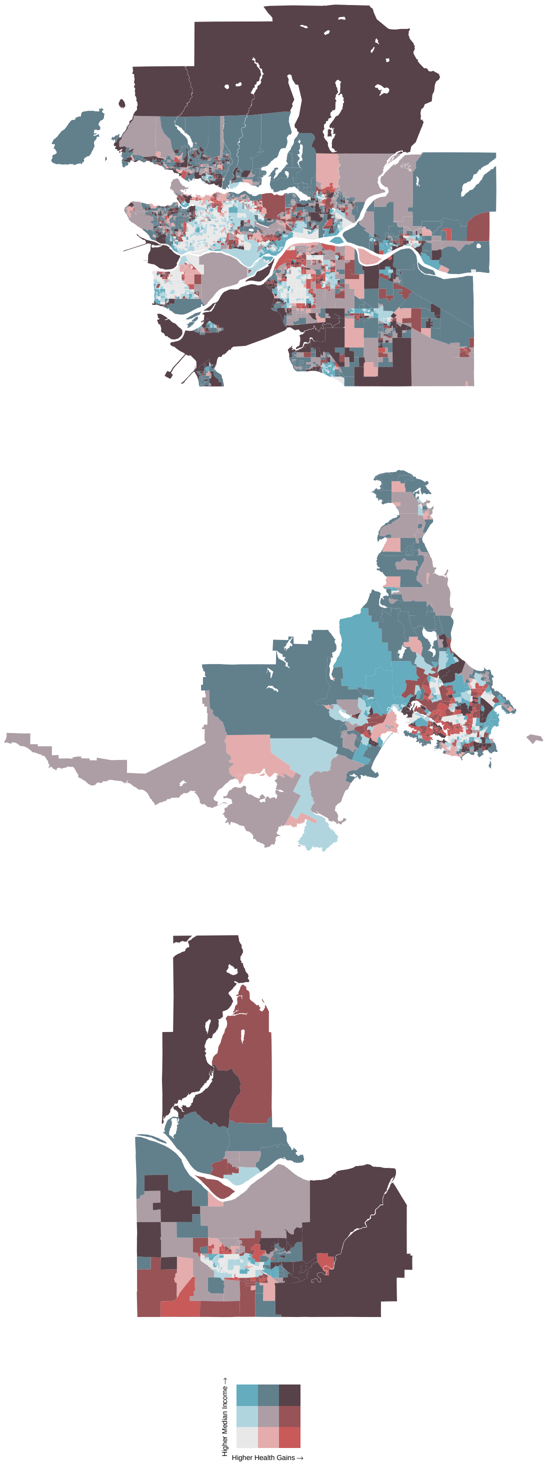

Figure F.2: Bivariate distribution of health gains using the RR from Lepeule *et al.* (2012) and median income for the Vancouver (top panel), Victoria (middle panel) and Abbotsford (bottom panel) CMAs.