



# Does a Carbon Tax Reduce CO<sub>2</sub> Emissions? Evidence from British Columbia

Felix Pretis<sup>1,2</sup>

Accepted: 16 March 2022 / Published online: 13 April 2022  
© The Author(s) 2022

## Abstract

Using difference-in-differences, synthetic control, and introducing a new break-detection approach, I show that the introduction of North America's first major carbon tax has reduced transportation emissions but not 'yet' led to large statistically significant reductions in aggregate CO<sub>2</sub> emissions. Proposing a new method to assess policy based on breaks in difference-in-differences using machine learning, I demonstrate that neither carbon pricing nor trading schemes in other provinces are detected as large and statistically significant interventions. Instead, closures and efficiency-improvements in emission-intensive industries in untaxed provinces have reduced emissions. Overall, the results show that existing carbon taxes (and prices) are likely too low to be effective in the time frame since their introduction.

**Keywords** Carbon tax · CO<sub>2</sub> emissions · Regulation · Break detection

## 1 Introduction

Mitigation of future climate change requires stringent curtailment of CO<sub>2</sub> emissions and carbon taxes have emerged as a tool-of-choice in numerous policy proposals (including the EU Green Deal) to implement these reductions (Carattini et al. 2019). Despite emission-reductions being their primary aim, few of the existing carbon taxes have been evaluated for their effectiveness in reducing emissions at observed tax rates. Existing studies that do look at the tax impact of emissions predominantly focus on single sector emissions (such as transportation) rather than assessing the aggregate impact. However, the impact on *aggregate* emissions is an important metric since anthropogenic climate change is driven by cumulative CO<sub>2</sub> emissions, mitigation of which requires *aggregate* reductions.

Here I assess the effectiveness of a carbon tax in reducing emissions by studying the impacts of the introduction of the first major carbon tax in North America: the carbon tax in British Columbia (BC), Canada introduced in 2008. The tax has been widely hailed as a success—both in popular perception as well as in its impact on consumer demand for fossil

---

✉ Felix Pretis  
fpretis@uvic.ca

<sup>1</sup> Department of Economics, University of Victoria, Victoria, Canada

<sup>2</sup> Nuffield College, University of Oxford, Oxford, UK

fuels, natural gas, and lack of impact on economic output (Murray and Rivers 2015). Nevertheless, no formal empirical impact estimates assess whether the tax has been measurably effective in reducing aggregate CO<sub>2</sub> emissions since its introduction. Existing studies that find large effects of carbon taxes in other countries tend to be overly optimistic in their assessment by focusing on individual sectors alone. Unlike other carbon taxes, such as the Swedish tax which predominantly affects transportation (and appears to have led to significant transport emission reductions—see Andersson 2019), the BC carbon tax lends itself to an assessment of the impact of the introduction of a carbon tax on *aggregate* emissions due to its broad coverage including roughly 70% of greenhouse gas emissions (Murray and Rivers 2015) with comparatively few exemptions.<sup>1</sup>

I study the emission response to the introduction of a carbon tax using two established approaches based on econometric policy analysis and then introduce a new econometric method to detect unknown interventions based on the detection of structural breaks. First, I assess the impact of the tax by imposing the tax as a known policy intervention in a panel difference-in-differences specification using province-level aggregated data accounting for the small number of treated units using randomisation inference, as well by using spatially-gridded emissions at a 1-degree spatial resolution. Second, I study the impacts of the introduction of the tax using synthetic control methods. Third, due to the myriad of potential interventions affecting CO<sub>2</sub> emissions, I propose a new econometric approach that agnostically detects treatment interventions (such as the introduction of the carbon tax) based on the detection of breaks using machine learning in panel fixed effects difference-in-differences models.

Following the maxim<sup>2</sup> that when the title is a question, the answer is ‘no’, I find that the introduction of a carbon tax is imprecisely estimated and has not ‘yet’ had a large and statistically significant impact on *aggregate* CO<sub>2</sub> emissions. The qualifier ‘yet’ is justified because the tax has led to some significant sector-specific emission reductions. In particular, the tax is associated with an estimated long-run 19% decrease in transportation-related emissions including personal vehicle traffic (when using difference-in-differences), and an average reduction of 5% when using synthetic control. I estimate the (short-run) semi-elasticity of transportation emissions with respect to the carbon tax rate as roughly a 0.2% reduction in emissions per dollar increase in the carbon tax rate, with a long-run estimate of 0.9% per dollar (albeit not accounting for introduction effects). The semi-elasticity of *aggregate* emissions, however, cannot be accurately estimated due to the large uncertainty around the point estimate. The absence of a statistically-significant effect on aggregate emissions can be reconciled with sector-specific reductions by the fact that some sectors (such as stationary combustion including manufacturing and heat production) in the taxed regions saw a (statistically insignificant) rise in emissions relative to control regions.

The identifying assumption needed in the present analysis is that there was no concurrent shock to the carbon tax that affected BC emissions relative to the control provinces conditional on the control variables—naturally it will never be possible to obtain perfect identification of the impact of carbon taxes on a large (country or province level) scale. Though I would argue that the absence of perfect identification should not

---

<sup>1</sup> The BC carbon tax’s main exemptions only include fuel exported from the taxed region, fugitive emissions, fuel used by shipping and air-travel to or from the region, and greenhouse operations in agriculture.

<sup>2</sup> See Betteridge’s law of headlines.

prohibit<sup>3</sup> the study of carbon pricing given its importance and prevalence as a policy tool to mitigate climate change. While perhaps a political surprise at the time, the carbon tax was not randomly assigned to the province of BC. However, if the introduction of a tax already reflects a higher willingness to reduce carbon emissions in the province, then policy-evaluation estimates yield an upward-biased estimate of the magnitude of the tax impact on emissions. Further, the results are robust to varying specifications of the control group.

Notably, the absence of a statistically significant effect of the introduction of the carbon tax on aggregate emissions does not imply there has been no effect. This particularly applies to the BC carbon tax for which only few years of data post-tax introduction exist, raising concerns about statistical power to detect an effect (though comparable studies such as Andersson 2019 rely on similarly short samples). However, given the praise carbon taxes have received in both literature and policy, the absence of a statistically detectable effect in aggregate emissions is sobering. Any tax-induced aggregate emission reductions do not appear to be detectable in the data and thus are far from climate policy targeting net-zero emissions by mid-century. Regardless of whether the absence of a statistically-significant reduction stems from a short sample, low tax rate, or incomplete coverage, the resulting implication for climate change is the same—there has not been a substantial detectable change in aggregate emissions in response to the introduction of this tax.

Conventional difference-in-differences and synthetic control approaches test whether the tax (or comparable policy) was effective in reducing emissions, but do not tell us whether other (previously unknown) interventions had significant effects. I therefore introduce a break-detection approach using machine learning in fixed effects panel estimation that allows for the detection of treated units and treatment dates without prior knowledge of their occurrence. Break detection is commonly applied in time series econometrics to detect policy interventions (and recently also in regression discontinuity designs—see Porter and Yu 2015), however, time series applications generally do not consider control groups making it difficult to identify causal effects. The proposed approach here presents an agnostic way of identifying shocks that affected individual units in the panel while using untreated units as control. Intuitively, a standard difference-in-difference or synthetic control estimator will tell us whether a particular intervention mattered, while the proposed approach here tells us which—previously unknown—interventions were statistically significant and when they occurred. Specifically, I propose to use indicator saturation methods (Castle et al. 2015) to detect breaks in individual fixed effects in panel, which correspond to treatment interaction dummy variables. When applied to CO<sub>2</sub> emissions, the introduction of the BC carbon tax is detected without prior knowledge as a structural break in transportation emissions, however, neither the tax nor carbon pricing in other provinces is detected in *aggregate* emissions. Instead, break detection identifies sudden changes in aggregate emissions in untaxed provinces which can be attributed to the closure of emission-intensive industries as well as efficiency improvements in electricity-generation.

Overall, the absence of a tax effect on aggregate emissions and the lack of breaks coinciding with carbon-pricing policy interventions suggests that existing carbon prices are set too low to achieve significant aggregate emission reductions in the (albeit short) time frame since their introduction, consistent with evidence from cross-country carbon pricing and policy studies (see e.g. Metcalf and Stock 2020; Rafaty et al. 2020; Eskander and

<sup>3</sup> See Akerlof (2020) and Ruhm (2019) on a discussion of tackling broad questions with imperfect identification.

Fankhauser 2020) At current rates, a carbon tax alone is likely insufficient in inducing substantial aggregate emission reductions and must be supplemented with complementary technology-pushing policies, ideally targeting sensitive interventions points (Farmer et al. 2019).

## 1.1 Background: The BC Carbon Tax and Carbon Pricing

British Columbia introduced a carbon tax on CO<sub>2</sub> emissions in 2008 at \$10 CAD per tonne emitted. While not randomly assigned, the introduction was unexpected as the government had not campaigned on the introduction of a carbon tax. Additionally, the tax was introduced rapidly, first proposed in February 2008 and implemented by July. The tax rate was initially set to rise \$5 CAD per year, this increase occurred until 2012, from when onwards it remained at 30\$ until 2018 when it was raised again to \$35 CAD per tonne of CO<sub>2</sub> emitted. The tax is designed to be revenue-neutral, making it a popular scheme with the public (Harrison and Peet 2012), with tax revenue returned to the public through tax credits and a dividend.

BC was the only province in Canada operating under its own carbon tax scheme, though the Canadian government introduced a federal carbon tax in 2019 (which falls outside of the sample under analysis here), and two other provinces have implemented major formal measures to curb emissions that fall within the sample available here<sup>4</sup>: Alberta uses a carbon price for industrial emitters above an emissions threshold (introduced in 2007), and Quebec operates under a cap and trade system (introduced in 2013 which appears to have resulted in some emission reductions when measured at the plant level, see Hanoteau and Talbot 2019). These alternative schemes raise some questions on the optimal choice of control groups. Since the carbon pricing policies in Alberta and Quebec are not of the form of a carbon tax, the main results presented here therefore include Alberta and Quebec in the control sample. Robustness checks omitting these provinces from the control group show that the results are unaffected by the inclusion of these two regions.

### 1.1.1 The Effectiveness of Carbon Taxes in Reducing Emissions

Analyses of the effectiveness of carbon taxes and prices have predominantly concentrated on metrics other than CO<sub>2</sub> emissions themselves. Where studies have assessed the impact on emissions, they primarily concentrate on specific sub-sectors. Green (2021) provides a comprehensive meta analysis of carbon pricing on emissions, finding mixed evidence of their effectiveness. Metcalf and Stock (2020) study the effects of carbon pricing in the European Union finding marginal reductions in emissions. Rafaty et al. (2020) assess the effect of carbon prices (rather than the introduction of a tax) on emissions in a sample of OECD countries, finding that while carbon pricing reduced emissions growth, current carbon prices are too low to show substantial emission reductions. This matches the results obtained by Lin and Li (2011) who study the introduction of a carbon tax in countries in Northern Europe finding significant impacts on per capita growth in emission reductions in only a small subset of countries, and the findings by Best et al. (2020) who find some emission reduction in response to carbon pricing. Assessing a single sector on a local level, Davis and Kilian (2011) study the effects of US gasoline taxes on carbon emissions finding

<sup>4</sup> Ontario introduced a cap and trade system in 2016.

a significant tax effect on vehicle emissions. This effect of a significant reduction in transportation emissions is also identified by Andersson (2019) who studies the introduction of a carbon tax in Sweden using synthetic controls but does not consider aggregate emissions due to the many exemptions in emission-intense sectors.

### 1.1.2 The Effectiveness of the BC Carbon Tax Across Different Measures

Murray and Rivers (2015) provide an early review of the literature on the general impacts of the BC carbon tax, though the empirical papers surveyed therein do not formally test the effect of the tax on emissions. There exist, however, a range of papers studying the impacts on employment and fuel demand. Xiang and Lawley (2018) and Gulati and Gholami (2015) find significant reductions in residential and commercial natural gas demand following the introduction of the tax. Rivers and Schaufele (2015) as well as Lawley and Thivierge (2016) and Antweiler and Gulati (2016) show that the carbon tax lead to a reduction in gasoline demand. Gittens (2018) provides evidence of the channel of this effect through incomplete pass-through of the tax. Saberian (2018) studies unintended consequences of the carbon tax, finding that the introduction of the tax led consumers to substitute to diesel cars thus increasing local pollution. Yip (2018) investigates the impact on labour markets and finds a small increase in unemployment of less-educated workers. Yamazaki (2017) assesses the employment impacts, finding a positive effect, though Azevedo et al. (2018) also assess the employment impact of the tax finding heterogeneous responses across industries, but no effect on aggregate employment. These opposite effects can be reconciled by concerns about the parallel trend assumptions in Yamazaki (2017) raised by Azevedo et al. (2018). Comparing econometric estimates to those of a general equilibrium model Carbone et al. (2019) find that employment shifted from more carbon-intense to less carbon-intense sectors.<sup>5</sup>

An early report around the time of the introduction of the tax predicted a 5% reduction in aggregate emissions using numerical simulations (BC 2008) though the accuracy of this prediction has not been assessed. Elgie and McClay (2013) use data up until 2011 to compute means and construct the non-parametric difference-in-differences without reporting standard errors. Their comparison finds a 9% reduction in per capita emissions following the introduction of the tax, however, due to the lack of standard errors these results do not provide a formal test.

Ahmadi and Yamazaki (2020) study the emission impacts of the BC carbon tax in the manufacturing sector at a plant level, finding marginal emission reductions in response to the tax. The authors also raise concerns about analysing the BC carbon tax using differences-in differences since the introduction of the carbon tax coincides with the Great Recession. The concern is valid when not accounting for underlying economic conditions, however, when controlling for changes in GDP—as implemented here and in the wider literature (see Metcalf 2019; Lin and Li, 2011; Rafaty et al. 2020), this is less of a concern, Shortly after the first release of the present paper, Metcalf (2019) also reported results of the tax impacts on the BC carbon tax using difference-in-differences, finding (marginally) significant aggregate emission reductions. However, these apparent reductions are not robust to the inclusion of dynamics in the model, choosing a subset of the available years,

<sup>5</sup> Beyond BC, Metcalf and Stock (2020) find little impact of the introduction of carbon taxes on GDP or employment in a sample of European countries.

or controlling for the small number of individuals in the panel. A further discussion contrasting the results is provided in Sect. 3.1.2.

The present paper makes three contributions to the literature. First, using difference-in-differences and synthetic control methods I show that there is little evidence of aggregate reductions in CO<sub>2</sub> emissions in response to the introduction of a carbon tax. These results are robust to the choice of the control group, the specification of the standard errors, randomisation inference, using spatially-gridded or province level data, accounting for biases in dynamic panel estimates, or the use of synthetic control methods to address concerns on non-parallel trends. Second, I show that despite the lack of an aggregate effect, there are heterogeneous effects across sectors with some (such as transportation) exhibiting significant large reduction in emissions consistent with existing evidence on carbon taxes. In addition, I also provide sector-level estimates of the emission elasticities with respect to tax rate. Third, I introduce an econometric break-detection method to assess the presence of treatment interventions, allowing for an agnostic approach to discover when interventions took place, and who was affected by these interventions. This identifies the closure of emission-intense mills combined with efficiency-improvements in electricity generation as likely effective measures in reducing emissions in three provinces.

## 2 Data

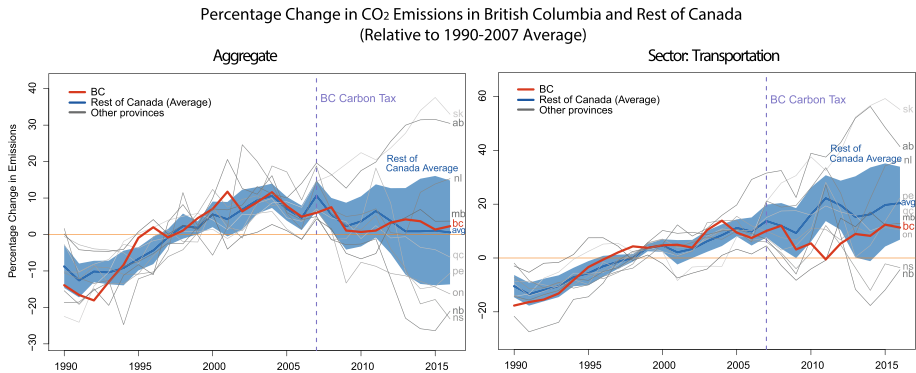
I obtain annual time series on provincial aggregate and sectoral CO<sub>2</sub> emissions from the National Inventory Report of Environment and Climate Change Canada (2018) from 1990 to 2016. Sector-level CO<sub>2</sub> emissions are categorised using the IPCC (Intergovernmental Panel on Climate Change) classification<sup>6</sup> and comprises stationary combustion sources including electricity generation; transportation including personal vehicles; fugitive emissions from coal, oil, and gas extraction; industrial processes including cement production; agriculture; and waste processing.

Provincial GDP (in constant 2007 dollars) and population data is available from Statistics Canada (2018) and included as socio-economic control variables to account for differences in population and economic development across provinces. Specifically, GDP allows for a control of differentiated impacts across provinces during the 2008/2009 recession.<sup>7</sup> Due to lack of available data, the Northwest Territories, Nunavut, and the Yukon are omitted from the sample, resulting in a (small) panel of N=10 Canadian provinces over T=27 years, one of which (BC) received the intervention of a carbon tax in 2008. The total regional sample spans 270 province-year observations. The annual regional emissions series are shown in percentage deviations relative to their 1990–2007 mean in Fig. 1 (for aggregate & transportation emissions—the single largest emitting sector in the taxed region) and for all available sectors in Fig. 2, together with the sectoral share of emissions over time for the taxed region. Summary statistics are shown in Table 1 for BC and provinces included in the control group.

Naturally working with a single treated province and a small number of post-treatment observations raises concerns about inference and statistical power. In particular, there are

<sup>6</sup> IPCC emission categories measure CO<sub>2</sub> emissions only. An earlier version of this paper used Canadian economic sector emission definitions which rely on CO<sub>2</sub> ‘equivalent’ emissions.

<sup>7</sup> Recent work by Metcalf and Stock (2020) shows little evidence of carbon taxes on GDP, thus supporting the inclusion of GDP as a control variable that itself is unaffected by the introduction of carbon taxes.

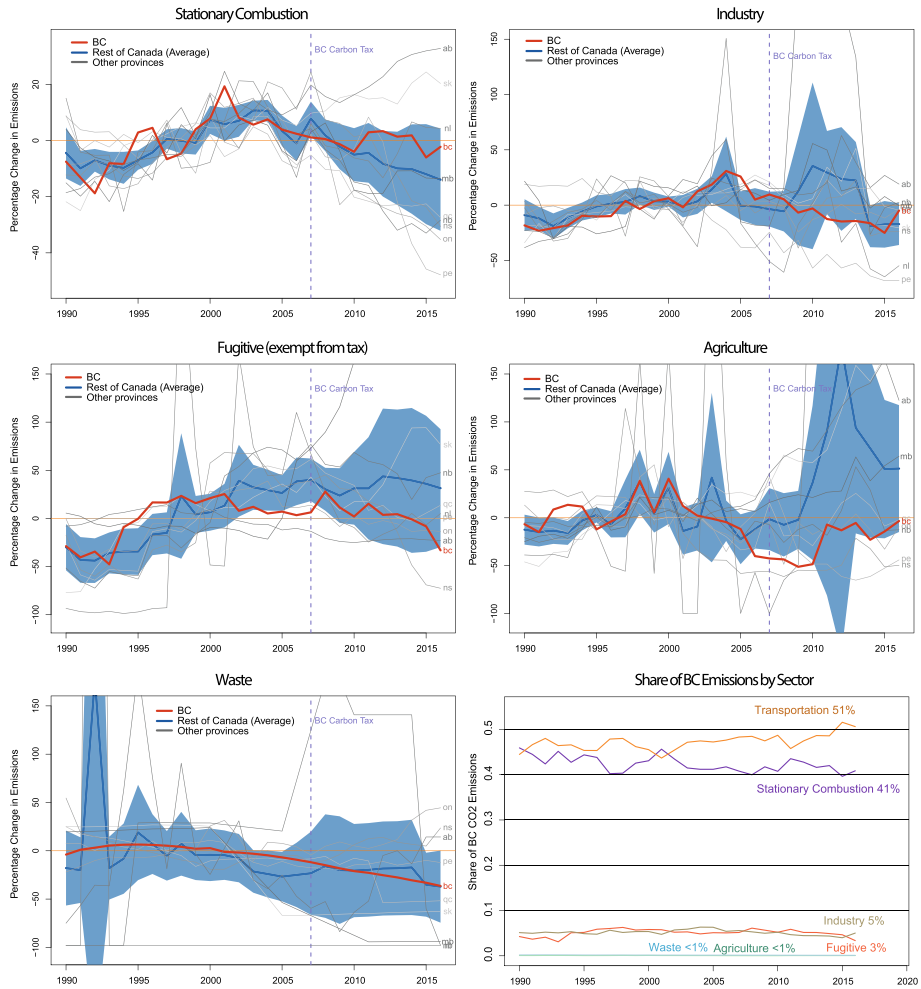


**Fig. 1** Percentage change in province-level aggregate CO<sub>2</sub> emissions relative to the 1990–2007 average for aggregate emissions (left) and transportation-emissions (right, the single largest emitting sector in the taxed region). Blue solid line corresponds to the average across included provinces other than British Columbia (BC) which is shown in red. The carbon tax is shown as purple dashed. Blue shading corresponds to the 95% confidence interval of the ‘rest of Canada’ average in each period [Canadian provinces are abbreviated as follows. British Columbia (bc), Alberta (ab), Ontario (on), Newfoundland and Labrador (nl), Nova Scotia (ns), New Brunswick (nb), Saskatchewan (sk), Quebec (qc), Prince Edward Island (pe), Manitoba (mb).]

concerns about inference based on difference-in-differences with a single treated region (Conley and Taber 2011). I use multiple approaches to address this in the difference-in-differences model, while also estimating the impact of the introduction of the carbon tax using synthetic control methods suitable for policy analysis with a single treated unit (Abadie and Gardeazabal 2003; Abadie et al. 2010). In the difference in differences setup, in addition to conventional estimates I report cluster-robust bootstrap standard errors (Cameron et al. 2008).<sup>8</sup> I also apply randomisation inference (see MacKinnon and Webb 2019) comparing the estimated tax coefficients to all possible coefficients when a hypothetical placebo tax is estimated at every other time point and province throughout the sample. To increase the number of available cross-sectional units and to assess the impacts of the tax on local emissions, I use spatially-gridded emissions data available from EDGAR (2017) from 1970 until 2012. To reduce concerns about local measurement errors I aggregated the gridded emissions from the available 0.1 degree spatial resolution to 1 degrees (approx. 111 km at the equator). This results in a total sample of 34,230 spatio-temporal observations spanning 815 grid-cells over 42 years of which 4,515 observations fall within the taxed region. The grid-level difference in CO<sub>2</sub> emissions is shown in Fig. 3. No annual GDP or population series are available on a gridded basis. Instead, I rely on lights-at-night as socio-economic controls in the gridded model (Zhang et al. 2016). Lights data is merged across 11 satellites by matching the means of overlapping years and averaged to 1-degree spatial resolution to match the emission data. Since my analysis focuses on Canadian provinces below 60°N, concerns about measurement above 60° (Feir et al. 2018) do not apply here.

Emissions are measured imprecisely which may result in substantial measurement errors in the dependent variable. Nevertheless, I do not expect measurement errors to present a major obstacle in the present analysis. First, measurement error in the dependent variable does not bias the coefficient estimates (though may amplify power concerns by increasing the noise). Second, I conduct the analysis using two different datasets to support the overall

<sup>8</sup> Clustered at the province level.



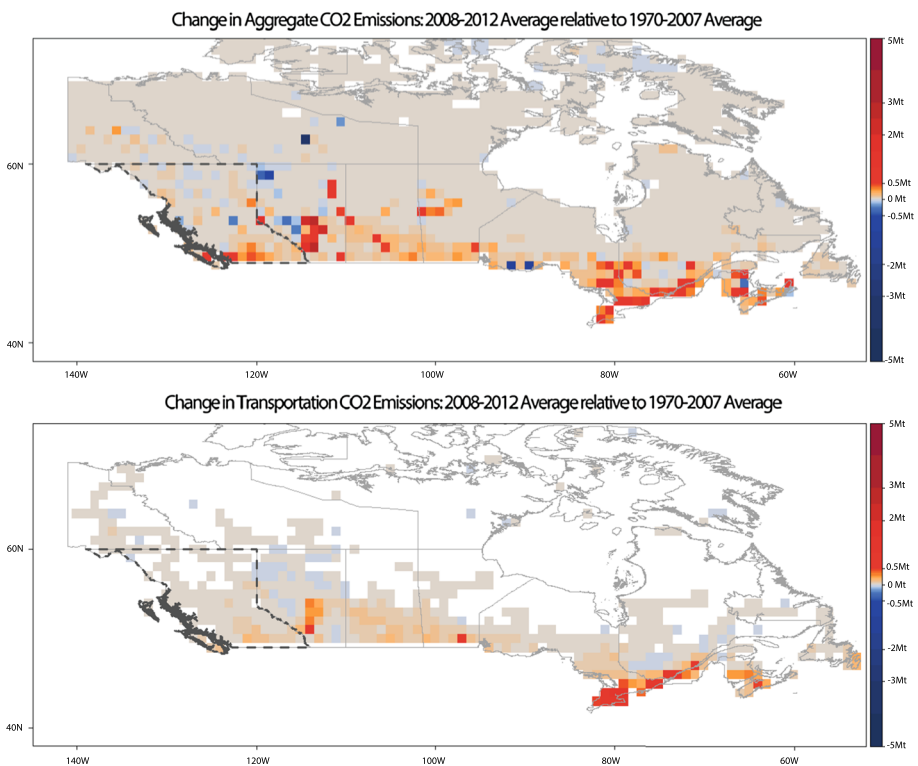
**Fig. 2** Percentage change in province-level sector-specific CO<sub>2</sub> emissions relative to the 1990–2007 average. Sector definitions correspond to IPCC emission categories. Blue solid line shows the average across included provinces other than British Columbia which is shown in red. Blue shading corresponds to the 95% confidence interval of the ‘rest of Canada’ average in each period. Lower right panel shows the share of emissions in the taxed province (BC) by sector over time

results. The first stemming from the national inventory reports for Canada, the second being the spatially-resolved EDGAR emission estimates with a standardised methodology across different countries. Even though EDGAR uses national records, the emission estimates are standardised across countries which may reduce some of the measurement error from national inventories. While some measurement error is unavoidable, an analysis of emissions rather than other measures of fossil fuel use is warranted since compliance with climate agreements (e.g. the Paris agreement) is primarily assessed on based on emissions reported in national inventories.

**Table 1** Province-level summary statistics for taxed region (BC) and control provinces 1990–2016

|                     | CO <sub>2</sub> emissions |              |         |       |        |       |        | Controls |        |
|---------------------|---------------------------|--------------|---------|-------|--------|-------|--------|----------|--------|
|                     | Total                     | Stat. energy | Transp. | Fugit | Indust | Agric | Waste  | Pop      | GDP    |
| <i>BC (taxed)</i>   |                           |              |         |       |        |       |        |          |        |
| Mean                | 47.17                     | 19.95        | 22.27   | 2.42  | 2.44   | 0.02  | 67.26  | 4.12     | 173.28 |
| SD                  | 3.78                      | 1.53         | 2       | 0.5   | 0.37   | 0.01  | 10.04  | 0.41     | 37.39  |
| Max                 | 52.22                     | 23.83        | 24.76   | 3.07  | 3.3    | 0.04  | 78.27  | 4.76     | 243.29 |
| Min                 | 38.29                     | 16.23        | 17.9    | 1.26  | 1.89   | 0.01  | 46.76  | 3.29     | 120.28 |
| <i>Control prov</i> |                           |              |         |       |        |       |        |          |        |
| Mean                | 54.72                     | 32.47        | 16.61   | 1.33  | 4.06   | 0.19  | 47.65  | 3.06     | 133.85 |
| SD                  | 62.85                     | 40.98        | 17.64   | 2.78  | 6.28   | 0.2   | 78.27  | 3.91     | 171.74 |
| Max                 | 211.66                    | 154.05       | 61.41   | 10.99 | 21.73  | 0.95  | 268.25 | 13.98    | 678.03 |
| Min                 | 1.18                      | 0.33         | 0.65    | 0     | 0      | 0     | 0.01   | 0.13     | 2.98   |

Emissions measured in Mt CO<sub>2</sub>; population in millions; GDP, billions in constant 2007 dollars. Control provinces here include: Alberta (ab), Ontario (on), Newfoundland and Labrador (nl), Nova Scotia (ns), New Brunswick (nb), Saskatchewan (sk), Quebec (qc), Prince Edward Island (pe), Manitoba (mb). Estimation results with different controls (omitting those that have carbon pricing or trading schemes are reported in Supplementary Material



**Fig. 3** Change in CO<sub>2</sub> emissions at a 1-degree spatial grid-level between the post-tax (2008–2012) and the pre-tax (1970–2007) average. Top panel shows the change in aggregate CO<sub>2</sub> emissions, lower panel shows the change in transportation CO<sub>2</sub> emissions. British Columbia (BC) which saw the introduction of a carbon tax in 2008 is shown with gray-dashed outline

### 3 Estimating the Emissions Impact of the BC Carbon Tax

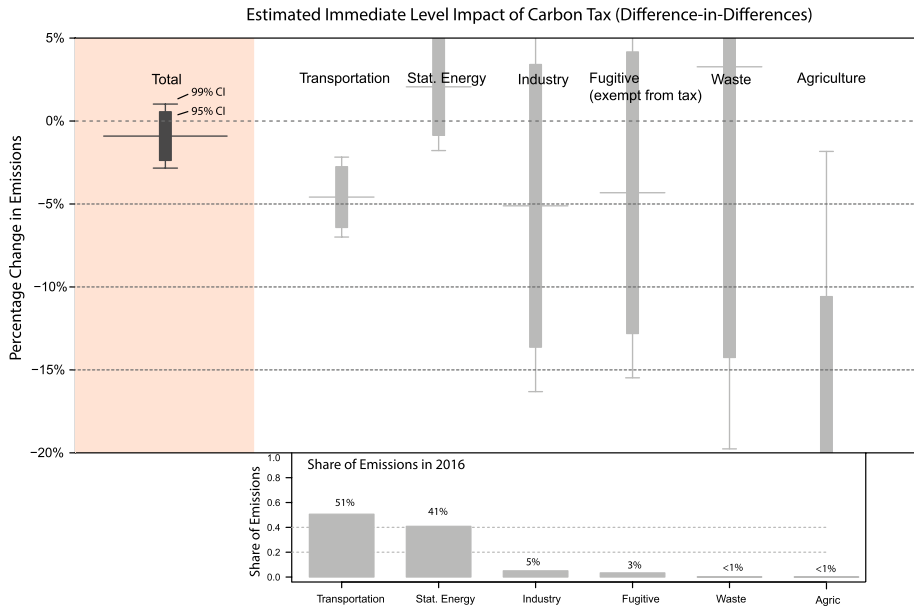
#### 3.1 Difference-in-Differences

To study the impact of the carbon on CO<sub>2</sub> emissions I estimate the difference-in-difference panel fixed effects model (1) and test the coefficient  $\beta$  of the variable of interest,  $Tax_{i,t}$ , which is a dummy variable equal to one for BC following the introduction of the carbon tax in British Columbia (in 2008), and zero otherwise. The per-tonne carbon tax has increased since its introduction; thus the single dummy variable captures the average impact of the introduction of the carbon tax. Section 3.1.4 also considers the semi-elasticity of emissions with respect to the tax by specifying  $Tax_{i,t}$  in (1) as the tax rate per tonne emitted. Modelling the incremental price increase using multiple intervention dummies—while feasible—would face even lower power than the present analysis due to the short time-span of available observations. Additional control variables in (1) include population, real GDP, and lagged emissions (to address potential autocorrelation), together with year-fixed-effects common across all provinces  $\lambda_t$  and province-fixed effects  $\mu_i$ :

$$\log(Emiss)_{i,t} = \lambda_t + \mu_i + \beta Tax_{i,t} + \alpha_1 \log(Emiss)_{i,t-1} + \gamma_1 \log(Pop)_{i,t} + \gamma_2 \log(GDP)_{i,t} + \epsilon_{i,t} \quad (1)$$

Due to the inclusion of lagged emissions, the parameter  $\beta$  captures the contemporaneous impact of the tax. To compute the estimated long-run effect of the tax I consider the equilibrium tax effect given by  $\beta_{Eqbm} = \beta / (1 - \alpha_1)$ . Standard errors of the estimate of  $\beta_{Eqbm}$  can be obtained using the delta method. The inclusion of an autoregressive term models persistence in emissions, not forcing any tax effect to have an immediate effect on the level of the dependent variable. Formal tests for autoregressive dynamics support the inclusion of one lag in (1)—diagnostic tests suggest little evidence of residual autocorrelation and longer lags are not statistically significant in most specifications and do not affect the overall results (see Appendix A). Additionally, the dynamic specification in (1) allows for tax effects on the growth rate instead of the level: if the tax effect predominantly affects the growth rate rather than the level, we expect the coefficient on lagged emissions to be close to one (see Appendix A for additional discussion). Difference-in-differences relies on parallel trends in the treatment and control groups. While parallel trends are not formally testable, it is common to estimate the model (1) including leads of the treatment dummy to assess whether parallel trends are rejected in the pre-treatment periods. In the present context with a single treated province, this amounts to a test of a single period indicator (zero–one dummy variable for a single period and province) in the pre-treatment periods. Single indicators can be estimated without bias, however, estimates will not be consistent (Hendry and Santos 2005). Visual inspection of the data suggests no obvious pre-trends (Fig. 1), formal tests for pre-trends based on leads of the tax treatment potentially identify one period ahead of the tax as an outlier but there is no obvious evidence of overall pre-trends (Appendix B). For robustness, Sect. 3.2 reports the results when using synthetic control methods (where results are consistent with findings when using difference-in-differences), and Sect. 3.3 considers break detection using an estimator robust to outliers (Fig. 4).

Earlier empirical evidence suggests significant decline in gasoline and residential natural gas demand in response to the introduction of the carbon tax. To assess the emission impact on a sectoral level, beyond aggregate effects I estimate model (1) for each



**Fig. 4** Top panel: estimation results—immediate impact estimates ( $\hat{\beta}$  in Eq. (1)) expressed in exact percentage change (standard errors clustered at province level transformed using the delta method). Lower panel: share of emissions by sector in 2016

sector defined by the IPCC using disaggregated CO<sub>2</sub> emissions. All panel models are estimated using the *plm* and *lfe* packages in R (Croissant and Millo 2008; Gaure 2013).

### 3.1.1 Aggregate Difference-in-Differences Estimation Results

The introduction of the BC carbon tax is imprecisely estimated but does not appear to have led to a detectable statistically significant reduction in aggregate CO<sub>2</sub> emissions. While the point-estimate of the introduction of the carbon tax on log aggregate CO<sub>2</sub> emissions is negative ( $-0.009$ ,  $se=0.008$  corresponding to a  $-0.09\%$  change, with a long-run estimate of  $-5\%$  with  $se=4\%$ ),<sup>9</sup> the coefficient is not statistically different from zero (Table 2, left column and Fig. 4). These results are unaffected by the exclusion of provinces that themselves have carbon-pricing schemes from the control group (Alberta and Quebec, see Table A1 in Appendix C), as well as accounting for potential biases in dynamic panel estimates (Nickell 1981)—I report Arellano-Bond (1991) GMM estimates of the dynamic panel in Appendix A.

The main results of the difference-in-differences estimation in Table 2 (and Fig. 4) are reported using standard errors clustered at the province level. Even though there are concerns with cluster standard errors when the number of clusters is small, the results here are robust to specifying conventional, clustered, or cluster-robust bootstrap standard errors (Cameron et al. 2008)—see Appendix D. To further alleviate concerns about the

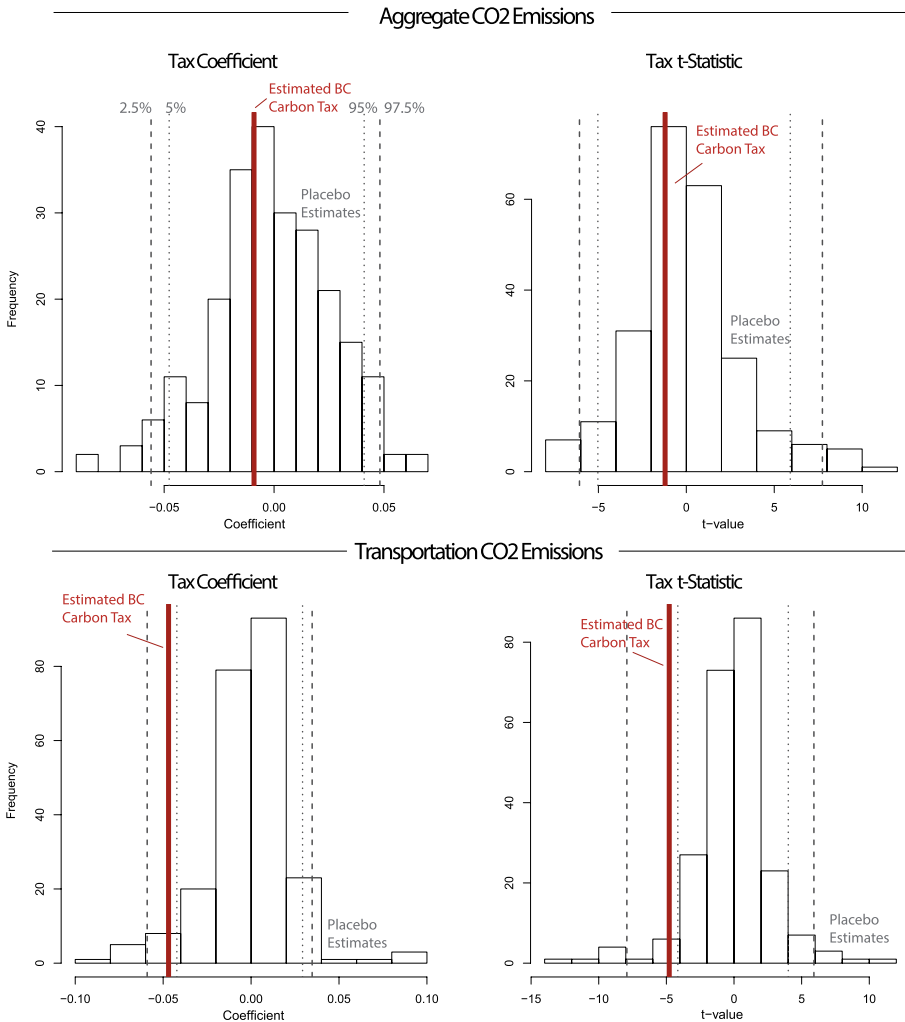
<sup>9</sup> The long-run estimate is a  $-5\%$  change with a standard error of  $4\%$ .

**Table 2** The effect of the carbon tax on CO<sub>2</sub> emissions (difference-in-differences estimation)

| Dependent variable: $\log(\text{Emissions})_{i,t}$ |                     |                     |                       |                     |                     |                     |                                 |
|--|---------------------|---------------------|-----------------------|---------------------|---------------------|---------------------|---------------------------------|
|  | Total               | Stat. Energy        | Transportation        | Fugit. (exempt)     | Industry            | Agriculture         | Waste                           |
| Carbon Tax <sub>i,t</sub>                          | - 0.009<br>(0.008)  | 0.020<br>(0.015)    | - 0.047***<br>(0.010) | - 0.044<br>(0.045)  | - 0.052<br>(0.046)  | - 0.485*<br>(0.231) | 0.032<br>(0.087)                |
| $\log(\text{Emiss})_{i,t-1}$                       | 0.809***<br>(0.051) | 0.834***<br>(0.053) | 0.778***<br>(0.060)   | 0.745***<br>(0.100) | 0.680***<br>(0.093) | 0.318***<br>(0.064) | 0.651***<br>(0.178)             |
| $\log(\text{Pop})_{i,t}$                           | - 0.068<br>(0.058)  | - 0.041<br>(0.111)  | - 0.103**<br>(0.035)  | - 1.566<br>(1.223)  | 0.464<br>(0.303)    | 0.475<br>(0.522)    | 0.421<br>(0.472)                |
| $\log(\text{GDP})_{i,t}$                           | 0.282***<br>(0.051) | 0.263***<br>(0.067) | 0.405***<br>(0.131)   | 0.902<br>(0.668)    | - 0.132<br>(0.317)  | 0.376<br>(1.108)    | - 0.028<br>(0.777)              |
| Eqbm. Tax Coef                                     | - 0.048<br>(0.044)  | 0.123<br>(0.093)    | - 0.211***<br>(0.038) | - 0.173<br>(0.214)  | - 0.164<br>(0.131)  | - 0.711*<br>(0.28)  | Pct. Impact<br>0.092<br>(0.252) |
| Eqbm. Se   | - 1%<br>(1%)        | 2%<br>(1%)          | - 5%***<br>(1%)       | - 4%<br>(4%)        | - 5%<br>(4%)        | - 38%***<br>(14%)   | 3%<br>(9%)                      |
| Pct. Impact Se                                     | - 5%<br>(4%)        | 13%<br>(11%)        | - 19%***<br>(3%)      | - 16%<br>(18%)      | - 15%<br>(11%)      | - 51%***<br>(14%)   | 10%<br>(28%)                    |
| Pct. Eqbm. Impact Se(4%)                           | 260                 | 260                 | 260                   | 260                 | 260                 | 260                 | 260                             |
| Observations                                       | 260                 | 260                 | 260                   | 260                 | 260                 | 260                 | 260                             |
| R <sup>2</sup> (within)                            | 0.717               | 0.693               | 0.768                 | 0.690               | 0.497               | 0.113               | 0.427                           |
| F Statistic (df=4; 221)                            | 140.04***           | 124.93***           | 182.49***             | 122.93***           | 54.54***            | 7.05***             | 41.21***                        |

Equilibrium impact corresponds to the contemporaneous tax impact scaled by the autoregressive coefficient. The estimated percentage (Pct) change for the dummy variable moving from 0 to 1 is given by  $100 \times (\exp(\hat{\beta}) - 1)$ . Standard errors shown in parentheses, obtained using the delta method for the equilibrium impact and percentage impact

\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001



**Fig. 5** Randomisation Inference—top panel shows results for aggregate emissions. Bottom panel shows results for transportation emissions. Left panel shows the distribution of placebo treatment estimates together with the actual treatment estimate (red). Right panel shows the distribution of placebo t-statistics, together with the actual treatment t-statistic. Dashed lines denote the 2.5th and 97.5th percentiles. Dotted lines denote the 5th and 95th percentiles. (Color figure online)

construction of standard errors with few treated units, I also use randomisation inference (see e.g. MacKinnon and Webb 2019; Buchmueller et al. 2011 on permutation tests). I estimate the above model (1) where the tax variable is modified to consider every time period in every control province as a placebo treatment. This results in 270 estimates of treatment coefficients and their t-statistics. I then compare the estimate of the actual BC carbon tax treatment effect (coefficient and t-statistic) to the distribution of estimated placebo tax coefficients and t-statistics and compute the resulting p-values. The results are robust to using randomisation inference (see Fig. 5). Comparing the actual carbon tax treatment estimates

to the placebo coefficients and t-statistics shows that the estimated aggregate tax effect and resulting t-statistics do not appear uncommon when compared to all placebo treatments. In other words, estimating a hypothetical treatment on provinces that were not subject to the tax results in estimated coefficients that commonly show a stronger response to the placebo treatment than the actual estimated treatment effect.

### 3.1.2 Reconciling the Results with Metcalf (2019)

Metcalf (2019) estimates a static model of log CO<sub>2</sub> emissions in a difference-in-differences fixed effects panel and finding marginally significant reductions in emissions as a result of tax. These results, however, are not robust and can be directly reconciled with the insignificant findings here: they rely on omitting controls for serial correlation in the data and not accounting for the small number of clusters when constructing standard errors. Either of these controls results in insignificant effects of the tax on aggregate emissions (see Appendix E).

### 3.1.3 Sectoral Analysis

How can the absence of a tax effect on aggregate emissions be reconciled with existing evidence of significant decline in BC gasoline demand? To answer this question, I turn to sector-level emission data which disaggregates CO<sub>2</sub> emissions to transportation (Transp.); stationary combustion sources (statEnergy); fugitive emissions from coal, oil and gas extraction (Fugit.); industrial processes (Indust) agriculture (Agric.); and waste processing (Waste). Emission series are shown together in Fig. 3. Fugitive emissions (which only make up a small fraction of overall emissions, approx. 3% in 2016) here act as a placebo test as they are entirely exempt from the tax. Greenhouse gas operations in agriculture are exempt, as well as non-fossil fuel greenhouse gas emissions, such as industrial processes and landfills, thus we expect small effects (if any at all) in the agricultural and industry sector.

A formal test for poolability confirms the heterogeneity of coefficients across sectors. I estimate a panel across sectors including sector-fixed effects and sector-year fixed effects together with interaction terms of the carbon tax (as well as each socio-economic control) with 6 sector dummy variables. An F-test for the joint significance of interaction terms strongly rejects the null hypothesis of the model being pool-able across sectors ( $p < 0.001$ ). This supports the analysis on a sector-by-sector basis. In particular, not only does the tax effect vary across sectors, but so do the relationships between sector-specific emissions and socio-economic controls. I therefore estimate the difference-in-differences estimation (1) for each sector with the results reported in Table 2 (right columns) and Fig. 4. Notably, population and GDP are insignificant in many of the sectors suggesting a potential decoupling of emissions from population and economic trends in Canada differing from results found in global studies (see e.g. Best et al. 2020).

There is evidence that the introduction of the tax has led to significant reductions in emissions in at least one out of the 6 sectors. Specifically, there is a significant decline in emissions from transportation (including vehicle traffic)—the single largest source of BC emissions (making up roughly 51% of emissions in 2016, see Fig. 4 lower panel). The estimated effect of the carbon tax is a reduction in transportation emissions of  $-5\%$  ( $se = 1\%$ )

in the short run, with an estimated equilibrium effect of  $-19\%$  ( $se=3\%$ ).<sup>10</sup> The significant reductions in transportation emissions are further supported when using randomisation instead of conventional inference (see Fig. 5). The estimated tax impact in BC lies below the majority of all possible placebo treatment estimates.

The magnitude of transportation emission reductions in BC are comparable to the 6–9% reduction in transportation emissions in Sweden following the introduction of a carbon tax (Andersson 2019). This magnitude of the effect perhaps is surprising given that the Swedish carbon tax (at roughly \$130 USD) is more than four times the current BC carbon tax at \$35 CAD, and was close to double the BC rate for the first years following its introduction. This may suggest that the emission semi-elasticity with respect to carbon taxes is higher in BC than Sweden, though a formal comparison cannot be made as Andersson (2019) does not provide estimates of the emission-elasticity with respect to the tax (for BC emission semi-elasticities see Sect. 3.1.4).

The difference-in-differences estimation also suggests a reduction of agricultural emissions, however, these impacts are more uncertain and not robust across methods. Specifically, the reduction in agriculture emissions is not statistically significant when using bootstrap standard errors (see Appendix D), randomisation inference (see Appendix F), or removing provinces that operate under carbon prices from the control group (see Appendix C). The point estimate of the carbon tax in industry is negative, consistent with the plant-level findings of Ahmadi and Yamazaki (2020), but imprecisely estimated here at the province level.

Why are the apparent sectoral reductions in transportation (and perhaps agriculture) not detectable in aggregate emissions? First, some sectors (such as stationary combustion) have seen an increase in emissions (albeit not statistically significant) relative to the control provinces. This is also apparent in the synthetic control analysis which shows an increase in emissions of observed relative to synthetic BC. Thus, aggregate emissions appear unchanged relative to the control group. Second, the sample of treated post-tax observations is small. There are only 9 years of emissions data available following the introduction of the tax, thus any aggregate effect may simply be drowned out by the variability of emissions.

### 3.1.4 The Semi-elasticity of Emission with Respect to the Tax Rate

The above difference-in-differences specification models the introduction of the tax using a single dummy variable. Here I report estimates of the response of emissions to the rate of the tax by estimating model (1) using the tax rate per tonne emitted. Consistent with above results, the uncertainty around aggregate emissions is too high to confidently provide an estimate of the semi-elasticity of aggregate emissions with respect to the tax. Notably, the estimation results here do not differentiate between an introduction effect and a price effect (see Rafaty et al. 2020).

The emission semi-elasticity (combining introduction and price effects—see Rafaty et al. 2020) for transportation emissions is statistically significant and estimated at an approximate 0.2% ( $se=0.1\%$ ) reduction in emissions per dollar increase in the tax rate (Table 3), with a long-run estimate of  $-0.9\%$  per dollar. This estimate captures the net

<sup>10</sup> Given the high level of persistence of transportation emissions (the estimated autoregressive coefficient is 0.78), 90% of the equilibrium effect would only be reached after approx. 10 years.

**Table 3** Emission elasticity with respect to the carbon tax—(using the tax rate)

|                                      | Dependent variable: $\log(\text{Emissions})_{i,t}$ |                     |                     |                     |                     |                     |                     |
|--------------------------------------|--|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                      | Total  | statEnergy          | Transp              | Fugit (exempt)      | Indust              | Agric               | Waste               |
| Carbon Tax<br>(\$CAD/tonne) $_{i,t}$ | - 0.0004<br>(0.0003)                               | 0.001<br>(0.001)    | - 0.002*<br>(0.001) | - 0.003<br>(0.006)  | - 0.001<br>(0.003)  | - 0.016<br>(0.014)  | 0.001<br>(0.011)    |
| $\log(\text{Emiss})_{i,t-1}$         | 0.809***<br>(0.051)                                | 0.833***<br>(0.042) | 0.776***<br>(0.043) | 0.744***<br>(0.049) | 0.680***<br>(0.049) | 0.320***<br>(0.065) | 0.651***<br>(0.051) |
| $\log(\text{Pop})_{i,t}$             | - 0.069<br>(0.057)                                 | - 0.039<br>(0.104)  | - 0.107<br>(0.068)  | - 1.582*<br>(0.662) | 0.458<br>(0.306)    | 0.426<br>(1.294)    | 0.423<br>(1.014)    |
| $\log(\text{GDP})_{i,t}$             | 0.285***<br>(0.051)                                | 0.263*<br>(0.127)   | 0.412***<br>(0.089) | 0.953<br>(0.690)    | - 0.145<br>(0.364)  | 0.394<br>(1.554)    | - 0.037<br>(1.230)  |
| Observations                         | 260  | 260                 | 260                 | 260                 | 260                 | 260                 | 260                 |
| R <sup>2</sup> (within)              | 0.717  | 0.693               | 0.767               | 0.690               | 0.496               | 0.112               | 0.427               |
| F statistic (df = 4;<br>221)         | 140.13***  | 124.84***           | 182.10***           | 123.13***           | 54.41***            | 6.94***             | 41.22***            |

\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

effect of both the introduction as well as the marginal increases of the tax rate on emissions themselves, implicitly accounting for the pass-through of the tax (such as through gasoline prices for the case of transport) and subsequently the response of those emitting CO<sub>2</sub> (such as consumers purchasing gasoline). Using the estimates in Table 3 allows us to quantify the net impact of the tax on emissions. For example, the \$5 increase in the per-tonne carbon tax in 2018 corresponds to an approximate \$0.02 increase in the retail gasoline price (from 6.7 to 8.9 cents per litre). Using the estimates from Lawley and Thivierge (2016),<sup>11</sup> this should result in a predicted 3.2% reduction in gasoline demand, which is of a similar order of magnitude as the estimated emission impact using estimates from Table 3, predicting a 1% ( $\$5 \times 0.002$ ) contemporaneous reduction in transportation emissions, with a long-run reduction of approximately - 4%.<sup>12</sup> The estimated long-run impact on transportation emissions is thus close to the estimated direct impact on gasoline fuel consumption.

### 3.1.5 Local Emissions Using Spatially-Gridded Data

Province-level aggregation of emissions results in a small number of observations in the sample. To increase the number of available cross-sectional units and also identify local impacts I use spatially-gridded emissions data. Unfortunately, no consistent spatially-gridded annual GDP or population controls exist for Canada for the time period matching the emission series. However, lights at night data is available and have been used in past geo-spatial analyses as proxies for local population and economic activity (see e.g. Donaldson and Storeygard 2016). Thus, I estimate panel models (1) using gridded emissions with lights at night as socio-economic control, including lagged autoregressive terms, grid-cell

<sup>11</sup> Lawley and Thivierge (2016) estimate the semi-elasticity of gasoline consumption as an approximate 1.6% reduction for a \$0.01 tax-induced increase in gasoline prices.

<sup>12</sup> The estimated long-run effect is roughly  $(-0.002)/(1-0.776) \times 5 = -4\%$ .

and year-fixed effects, as well as the tax intervention. Standard errors are clustered at the province (treatment) level (using gridded observations results in a large number of individual grid cells in each cluster). As an additional robustness check (and given the estimation results strongly supporting growth effects) I also report the regression results when modeling the first difference of log emissions.

To assess locally-varying impacts of the tax, I interact the tax intervention dummy with a full set of individual fixed effects for the treatment region (each grid-cell within BC is allowed to deviate from the common tax coefficient). I assess both individual significance of the interaction terms (to identify regional variation in tax impacts) as well as overall significance of the global hypothesis of no tax impact using the Simes (1986) modified Bonferroni correction for multiple testing. Future work may explore a spatial regression discontinuity design across the borders of the taxed region, though region-specific challenges (such as the Rocky Mountains) and low population density near the borders may reduce the informativeness of this approach.

Using spatially-gridded data confirms the absence of a significant tax effect on aggregate emissions. The coefficient on the tax intervention is not statistically different from zero in the gridded model (Table 4) of aggregate emissions while controlling for local socio-economic development using satellite-derived lights-at-night observations. Nevertheless, there is some evidence of a regional response to the tax. The tax effect is statistically significant and negative for aggregate emissions for some regions (Fig. 6) and we reject the global null hypothesis of no effect when correcting the grid-based tests for multiple testing (using Simes 1986). Some regions that systematically show a decline in emissions relative to others when allowing the effect of the tax to vary over space—a notable cluster is in the North East of the province which constitutes a large oil and gas field. The introduction of the carbon tax coincides with the ‘great recession’, so the estimated reduction in emissions following the tax may reflect downturn in a particular emission-intensive industry in the treated region that is independent of the tax. Using spatially-disaggregated transportation emissions also confirms the findings when using aggregate data. The tax has led to a significant reduction in transportation emissions, with the immediate impact estimate of a reduction in emissions of 1.3%. In levels, this effect is notably smaller than the estimate obtained when using province-level data (Sect. 3.1.1). This difference in magnitude can be driven by the different temporal coverage of the samples. The spatially-gridded sample only spans up until 2012 compared to 2016 for the province-level data. Additionally, the autoregressive coefficient in the spatially-gridded panel is close to one, supporting an analysis in first difference, and thus perhaps capturing an effect on the growth rate of transportation emissions.

### 3.2 Synthetic Control

To further alleviate concerns about difference-in-differences with a single treated unit, I investigate the effect of the introduction of the carbon tax using synthetic controls (Abadie and Gardeazabal 2003; Abadie et al. 2010). This approach is suitable for a single treated region and relaxes the (difficult to verify) parallel trend assumption underlying difference-in-differences by constructing a synthetic (estimated) version of the treated region, and subsequently compares the synthetic (untaxed) region to the observed region under the tax. The approach is appealing in the case of a single treated region, but the small number of control regions means that inference in synthetic control is less straight forward. I apply the

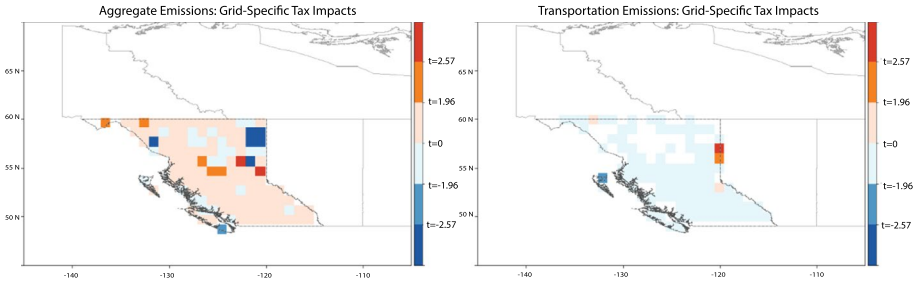
**Table 4** Gridded Emissions Estimates when estimating model (1) using spatially-gridded aggregate emissions (left) and transportation emissions (right)

|                                      | Aggregate CO <sub>2</sub> Emissions |                            | Transportation CO <sub>2</sub> Emissions |                           |                           |                            |                       |                       |
|--------------------------------------|-------------------------------------|----------------------------|--|---------------------------|---------------------------|----------------------------|-----------------------|-----------------------|
|                                      | log(Emiss) <sub>i,t</sub>           | Δlog(Emiss) <sub>i,t</sub> | log(Emiss) <sub>i,t</sub>                | log(Emiss) <sub>i,t</sub> | log(Emiss) <sub>i,t</sub> | Δlog(Emiss) <sub>i,t</sub> |                       |                       |
| Carbon Tax <sub>i,t</sub>            | 0.016<br>(0.039)                    | 0.005<br>(0.005)           | 0.009<br>(0.011)                         | 0.008<br>(0.011)          | 0.040<br>(0.026)          | - 0.013***<br>(0.002)      | - 0.007***<br>(0.001) | - 0.013***<br>(0.002) |
| log(Emiss) <sub>i,t-1</sub>          | 0.935***<br>(0.011)                 | 0.919***<br>(0.021)        | 0.919***<br>(0.021)                      |                           | 1.022***<br>(0.003)       | 0.994***<br>(0.0004)       |                       |                       |
| Light <sub>i,t</sub>                 |                                     | 0.00001<br>(0.00002)       |  |                           |                           | 0.00003***<br>(0.00001)    |                       |                       |
| ΔLight <sub>i,t</sub>                |                                     |                            |  | - 0.007*<br>(0.003)       |                           |                            |                       | - 0.002<br>(0.001)    |
| Observations                         | 72,505                              | 34,230                     | 31,728                                   | 30,200                    | 23,865                    | 11,655                     | 23,310                | 11,100                |
| R <sup>2</sup> (within)              | 0.99                                | 0.99                       | 0.99                                     | 0.166                     | 0.99                      | 0.99                       | 0.57                  | 0.32                  |
| F Statistic<br>(df=1-4;<br>554-1511) | 0.177                               | > 200***                   | > 500***                                 | 0.697                     | 2.483                     | > 200***                   | 46.56***              | 17.32***              |

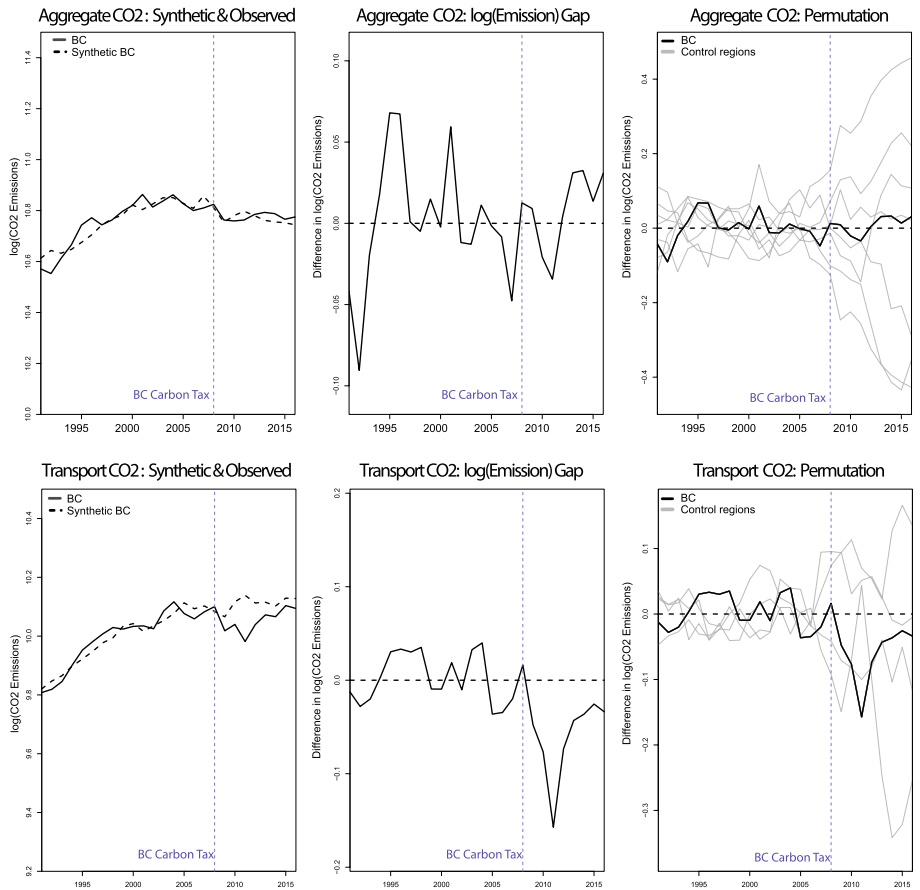
Robust standard errors shown in parentheses

Models estimated in levels and first differences (denoted by Δ)

\*\*\*p < 0.05, \*\*p < 0.01, \*p < 0.001



**Fig. 6** Individual Significance of Tax Effect: local grid-cell specific t-statistic of the tax within British Columbia when estimating model (1) using spatially-gridded observations (from Sect. 3.1.5) and grid-specific tax coefficients. Aggregate emissions shown in the left panel, transportation emissions shown in the right panel. Blue denotes significant negative tax effect on emissions, red indicates a significant positive tax effect on emissions. Light blue and light red/pink denotes insignificant negative/positive point estimates



**Fig. 7** Synthetic control analysis of the BC carbon tax for aggregate (top) and transport emissions. Left panels show synthetic and observed log CO<sub>2</sub> emissions in BC, middle panel shows the difference (gap) between log emissions of synthetic and observed BC. Right panels show the gap for the permutation across all control regions

permutation test (Abadie et al. 2010) where each control region is treated as a placebo and assigned a placebo intervention in 2008 when the carbon tax was introduced. Subsequently I compare the difference between observed BC and synthetic BC to the difference in every placebo region using the ratio of pre- and post-treatment mean square prediction errors. Nevertheless, the construction of p-values based on the observed deviations of the placebo-permutation study here is only marginally informative as only 9 untreated regions are available. In other words, the difference of observed and synthetic BC can only be compared to at most 9 placebo regions. Synthetic control is implemented here using the *Synth* package in R (Abadie et al. 2011).

Synthetic control results using province-level data are consistent with the estimates from the analysis using panel difference-in-differences (Fig. 7 and Appendix G—which also reports the weights on control provinces). There is little noticeable difference in aggregate emissions between BC and synthetic BC. The gap in emissions between synthetic and observed BC ranges from  $-3\%$  to  $+3\%$  (Fig. 7, top panel). Even though there are only a small number of control provinces and therefore the permutation test is of limited use, the results also support the absence of effects on aggregate emissions (Appendix G). The synthetic control analysis also confirms sector-level results of the difference-in-differences estimation. Transportation emissions in observed BC are around 3–15% lower than synthetic BC (Fig. 7, lower panel) with an average effect of  $-5\%$  over the treated period. The apparent re-bounce towards the end of the sample could be driven by the absence of a planned increase in the carbon tax from 2012 onwards. The original announcement of the tax aimed for a steady increase of the tax by \$5 dollars per tonne of CO<sub>2</sub> per year, however instead the tax was held constant from 2012 to 2018. Synthetic control also shows that agriculture is not well-approximated by the empirical model (see Appendix G, similar to difference-in-differences as evidenced by the high standard error when using bootstrap–cluster robust errors).

### 3.3 Break-Detection to Assess Policy Impacts

An alternative approach to imposing the policy intervention *a-priori* as in conventional difference-in-differences or synthetic control is to ask the data-driven question: have there been sudden changes in emissions, and if so, do these align with policy interventions? I propose to use break-detection methods to discover level-shifts in individual fixed effects  $\mu_i$  in a panel model to identify treatment time and treated units (Castle et al. 2015). Such level shifts are equivalent to treatment dummy variables in a difference-in-differences specification where the period before the shift is the pre-treatment period, the period after the shift is the post-treatment period, and the unit for which the shift has been detected corresponds to a treated unit. The retained step-shifts identify unit-specific heterogeneous treatment effects (see Pretis and Schwarz 2022 for more detail). Thus, this method allows treatment dates and units to be identified without prior knowledge of treatment timing or assignment.

While standard difference-in-differences approaches act as a test of whether a known intervention had an effect, using break detection in the similar context allows us to identify what mattered, and if any of the detected effects coincide with known interventions (such as the introduction of a carbon tax). Thus, difference-in-differences answers the question of whether a known policy mattered while break detection evaluation answers the question of what mattered (that might have been unknown previously). The challenge is then to attribute the detected breaks to interventions that occurred during the identified time periods.

Detecting breaks with subsequent policy (or event) attribution has become common practise in time series analysis (and recently in regression discontinuity designs, see e.g. Porter and Yu 2015, or Hansen 2017), however, has not been used in panels to detect treatment. In time series for example, in early work, Perron (1989) detects breaks in GNP time series attributable to the Great Depression and an oil price shocks. More recently, Hendry (2018) employs a break detection approach in a single time series to detect policy interventions in UK CO<sub>2</sub> emissions, Pretis et al. (2015) use break detection to evaluate the accuracy of climate models, and Mukanjari and Sterner (2018) use break detection as a modified univariate event-study method. Williams and Reade (2016) detect jumps in Twitter data aligned with news-worthy events, and Ericsson (2017) uses breaks to identify periods of varying forecast accuracy in forecasts of the US debt. Estrada et al. (2013) employ break detection to assess whether the Montreal Protocol significantly affected CFC emissions and subsequently temperatures, and Schneider et al. (2017) as well as Pretis et al. (2016), and Pretis (2020) use break detection to identify volcanic eruptions in temperature records. Focusing on policy interventions, Apergis and Lau (2015) assess whether breaks in Australian electricity markets can be attributed to introduction of climate policy. In a micro-econometric setting Piehl et al. (2003) use break detection to evaluate policy on youth homicide reduction. However, these existing time series applications of break detection to evaluate policy (and interventions in general) do not have control groups, thus making a causal interpretation difficult.

Here I generalise break detection to a fixed-effects panel setting. This is comparable to a difference-in-differences event study where every time period is considered a potential event and every unit is considered as potentially receiving treatment while un-treated units without breaks act as controls. The approach is also comparable to the regression discontinuity design in Porter and Yu (2015) where every level of the running variable may provide a discontinuity which is unknown *a-priori*. In follow-up work, we further show the equivalence between structural breaks and heterogeneous treatment effects in Pretis and Schwarz (2022) with applications to EU transport emissions in Koch et al. (2022).

### 3.3.1 Break Detection Approach to Detecting Treatment

The starting point for the proposed approach here is to specify a general (fully-saturated) model allowing for a treatment dummy for every time period and every unit. This corresponds to a model with more variables than observations, which can be reduced to a specific model using variable selection algorithms from machine learning (for example general-to-specific selection or shrinkage based estimators such as the LASSO, see Tibshirani 1996). The general model is given by Eq. (2):

$$\begin{aligned} \log(Emiss)_{i,t} = & \lambda_t + \mu_i + \beta Tax_{i,t} + \alpha_1 \log(Emiss)_{i,t-1} + \gamma_1 \log(Pop)_{i,t} \\ & + \gamma_2 \log(GDP)_{i,t} + \sum_{j=2}^T \delta_{i,j} 1_{\{t \geq j\}} + \epsilon_{i,t} \end{aligned} \quad (2)$$

where  $1_{\{t \geq j\}}$  denotes an indicator function for the interaction of each  $t$  and  $i$ , in other words the difference-in-differences estimator for each time period and unit. This allows for a step-function (capturing treatment) at time  $t=j$  for any unit  $i$  with heterogeneous treatment magnitude given by the estimated coefficients  $\delta_{i,j}$  at any point in time. As many units are likely 'untreated', the coefficient matrix of all  $\delta_{i,j}$  will be sparse (zero coefficients for most units and time periods) and we can rely on machine learning model selection methods for sparse models with more variables than observations. Here I use indicator saturation methods

(Castle et al. 2015 implemented in R using `gets`, see Pretis et al. 2018a, b), though the LASSO and related estimators are also feasible.

Using a general-to-specific indicator saturation approach, the general model (2) is reduced to a specific one (removing treatment dummy variables) using automated model selection at a chosen target significance level  $p$  that determines the expected number of false-positives (controlling for multiple testing) given by  $pN(T-1)$ , where  $NT$  denotes the panel sample size in a balanced panel.<sup>13</sup> Johansen and Nielsen (2009, 2016) provide the asymptotic theory for simple impulse indicators and Nielsen and Qian (2018) extend the results to step-indicators in time series, showing that the false-positive rate converges to the chosen target significance level  $p$ . Thus, the false-detection rate is easily controlled by altering  $p$ . For example, setting  $p=0.05$  yields an expected false-positive rate of 0.05 detected treatment interventions. The final model then includes a subset of  $k_i$  of all possible break variables. The detected break times here are denoted by  $T_{ij}$ :

$$\begin{aligned} \log(Emiss)_{i,t} = & \lambda_t + \mu_i + \beta Tax_{i,t} + \alpha_1 \log(Emiss)_{i,t-1} + \gamma_1 \log(Pop)_{i,t} \\ & + \gamma_2 \log(GDP)_{i,t} + \sum_{j=1}^{k_i} \delta_{i,j} 1_{\{t \geq T_{i,j}\}} + \epsilon_{i,t} \end{aligned} \quad (3)$$

Note that not all units may exhibit a significant break (i.e. treatment) at all, so  $k_i=0$  for most  $i$ . Intuitively this approach searches over zero–one dummy variables at every point in time for each  $i$ , generalised here to a panel setting where the retained indicator can be interpreted as treatment affecting unit  $i$ , in period  $T_{i,j}$ .

As an illustration, consider a province that implemented a policy in a particular year leading to a significant reduction in emissions relative to all other provinces. A researcher may, however, have not been aware of this policy and so is unable to estimate a conventional difference-in-differences or synthetic control model. The proposed approach assesses every period for every province as a potential treatment intervention, where significant interventions are retained and insignificant removed from the model (at a pre-specified level of significance  $p$  which determines the false-positive rate). Having identified individuals in which breaks occur together with the estimated break dates, break-detection requires ex-post attribution of detected interventions. If timing coincides with the previously unknown policy intervention, this can provide evidence for its effectiveness. Naturally, had this policy been common knowledge, then it could have been tested in a conventional difference-in-differences (or synthetic control) approach.

The proposed break detection method can be interpreted in two ways. First, as a robustness check for conventional difference-in-differences analysis where a known policy to be tested is `forced` in the model and not selected over, while selecting over potential unknown treatment. This allows the known policy to be evaluated while simultaneously testing whether there have been other significant interventions affecting the outcome variable since omitting relevant variables (such as previously unknown other policy interventions) may bias the estimates of the intervention being tested. This follows the theory-embedding approach of Hendry and Johansen (2015) where some variables are not selected over, while simultaneously selecting over a full set of break-variables, a subset of which can describe any sudden change that may have been unknown *a-priori*. Second, break detection can act as a purely data-driven method to detect significant interventions

<sup>13</sup> The maximum number of step-shifts in a balanced panel is  $N(T-1)$ , as for each  $N$  there are  $T-1$  possible breaks—a break in the first period is identical to the constant fixed effect, the first possible break occurs at  $t=2$ . Because of the presence of a lagged term, the number of available time periods is 26.

**Table 5** Break detection results for aggregate CO<sub>2</sub> emissions at a significance level of detection of  $p=0.025$ 

|                                  |                    | Dependent variable: $\log(\text{Emissions})_{i,t}$ |   |
|----------------------------------|--------------------|--|---|
|                                  |                    | Coefficients<br>(not selected-over)                |   |
| $\log(\text{Emiss})_{i,t-1}$     |                    | 0.74*** (0.035)                                    |   |
| $\log(\text{Pop})_{i,t}$         |                    | 0.02 (0.085)                                       |   |
| $\log(\text{GDP})_{i,t}$         |                    | 0.19* (0.088)                                      |   |
| Negative detected breaks         | Break coefficient  | Eqbm. Impact                                       | Attribution                               |
| New Brunsw. 2012 ( $\pm 1$ year) | - 0.083*** (0.02)  | - 0.319*** (0.081)                                 | Restructuring electricity-generation      |
| %Pct. Impact                     | - 8%*** (2%)       | - 27%*** (6%)                                      |   |
| Nfld & Lab. 1999 ( $\pm 0$ year) | - 0.132*** (0.032) | - 0.506*** (0.151)                                 | Hydro-Electric capacity improvement       |
| %Pct. Impact                     | - 12%*** (3%)      | - 39%*** (9%)                                      |   |
| Nova Scotia 2012 ( $\pm 2$ year) | - 0.066** (0.018)  | - 0.253** (0.076)                                  | Electricity restructuring & mill Shutdown |
| %pct. impact                     | - 6%*** (2%)       | - 22%*** (6%)                                      |   |

The introduction of the BC carbon tax, Alberta and Quebec carbon pricing schemes are not detected as significant interventions in aggregate emissions

Indicator saturation  $p\text{-alpha}=0.025$ . Outliers and positive breaks not shown—please see Appendix I. Approximate 95% confidence interval on break dates computed using Hendry and Pretis (2016)

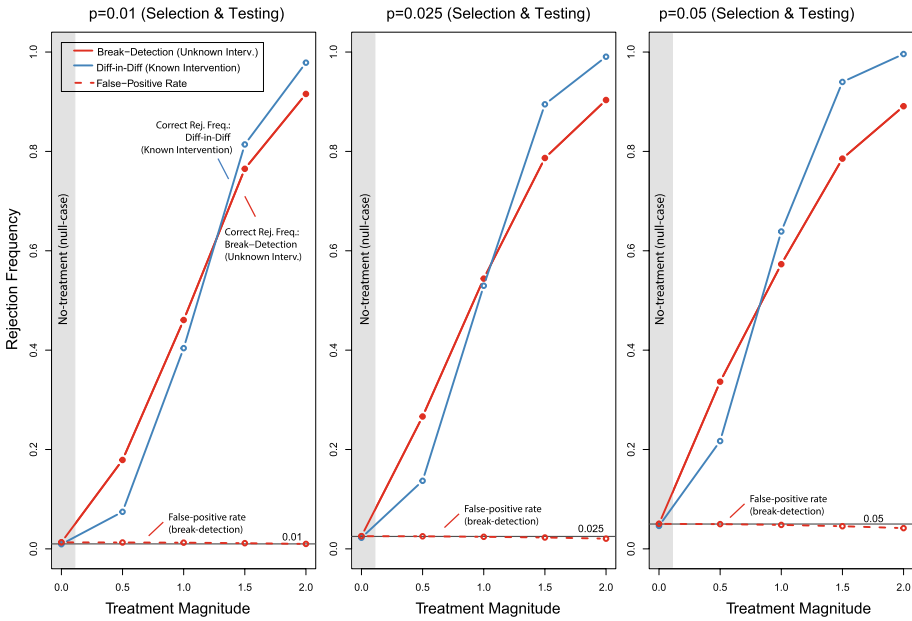
\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

without imposing any single intervention from the outset. Here this allows for detection of abnormal periods of emissions relative to the remaining control group.

Break detection using step-indicator functions has high power when the underlying data generating process is well-approximated by a step-shift change. However, some emission policies may have induced more gradual changes, such as a trend change in CO<sub>2</sub> emissions. These more gradual changes can be detected by either searching for breaks in the first difference of the series (where a step-shift in first differences corresponds to a trend break in the level). An alternative is to look for breaks in individual trend functions in the first place, as for example recently employed by Walker et al. (2019) who assess previously unknown changes in prescription patterns of medical doctors. For the present application I focus on step-shifts in the individual fixed effects, which risks missing some of the more gradual change (as perhaps induced in Ontario following the phase-out of coal in the 2000s).

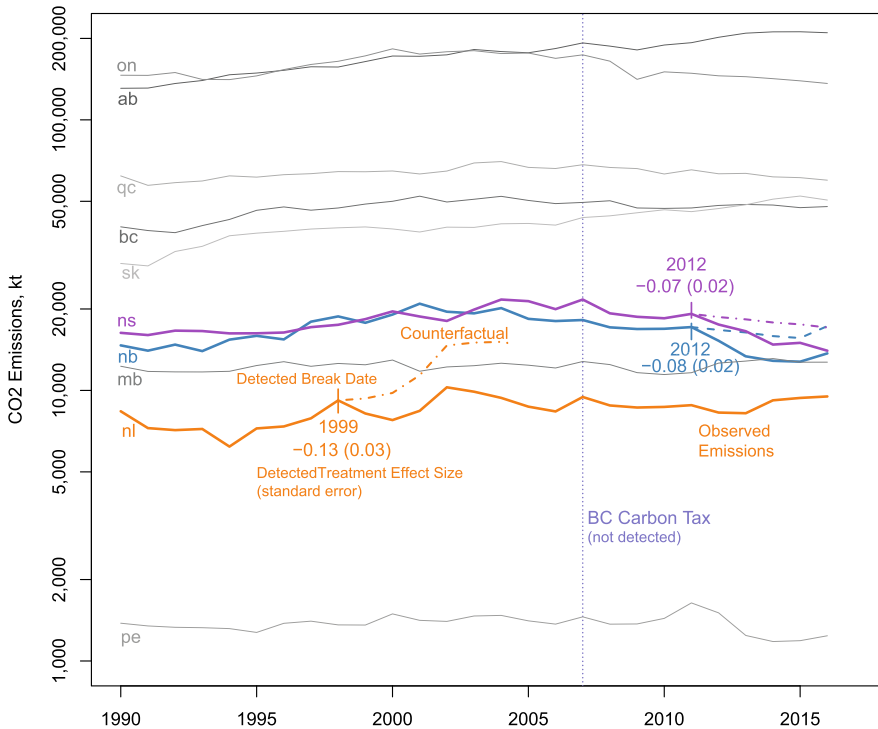
### 3.3.2 Break Detection: Illustrative Monte Carlo Simulation

To investigate the properties of the proposed break detection approach I consider a simple Monte Carlo exercise (details provided in Appendix H) that closely matches the sample of Canadian emission. I simulate a single treated unit with varying levels of treatment (in standard deviations of the error term) in a sample of  $T=27$ ,  $N=10$  observations. Simulation results in Fig. 8 show the detection frequency of the unknown intervention compared to the null rejection frequency when the intervention is known and imposed in a conventional difference-in-differences panel. The results demonstrate that



**Fig. 8** Simulation results comparing a known difference-in-differences intervention to detecting this intervention when it is unknown using break detection. The figure shows the detection/rejection frequency and false positive rate of break-detection (unknown intervention, red) and difference-in-differences (known intervention, blue) for different magnitudes of the treatment intervention in SD of the error term, including the no-treatment case (gray shaded). The correct rejection frequency (i.e. identification of treatment) in the unknown treatment break detection approach (red solid lines) is close to the optimal known treatment case (blue solid lines). The false positive rate of break detection is stable around the chosen level of significance (dashed lines). The false-positive rate of the break detection approach is shown as red-dashed and closely matches the nominal level of significance of selection ( $p=0.01$  left,  $p=0.025$  middle, and  $p=0.05$  right). Simulation results based on 5000 replications. Full results are also given in Appendix Table H1. (Color figure online)

treatment can be identified without prior knowledge of its occurrence at a rate that is close to the optimal case where the intervention is known and tested using difference-in-differences (see also Appendix H for more details). For example, consider a treatment intervention of 2SD of the error term in Fig. 8: the null-rejection frequency of the *known* difference-in-differences intervention dummy of 98% (at a 1% level of the test, blue in Fig. 8) is close to the detection frequency of 97% when the treatment is *unknown* (at a false-detection rate of 1%, red in Fig. 8). Simulation results also confirm that the false-positive rate of detected interventions is easily controlled by adjusting the level of significance of selection (dashed lines in Fig. 8). The simulations in Appendix H further show that the indicator saturation implementation of treatment detection outperforms the LASSO both in terms of power of detecting treatment, as well as in controlling the false-positive rate.

Detected Treatment Breaks in Agg. CO<sub>2</sub> Emissions (Difference-in-Differences Specification)

**Fig. 9** Detected negative breaks (interventions) in the panel fixed effects model of log aggregate emissions (3). Detected breaks (relative to the control group) are shown as dashed vertical coloured lines for the three provinces in which interventions are identified. Break magnitude is shown as the estimate immediate impact. Three negative breaks (reductions in emissions) are identified and shown in bold: New Brunswick in 2012 ( $\pm 1$  year 95%CI), Nova Scotia in 2012 ( $\pm 3$  year 95%CI), and Newfoundland and Labrador in 1999 (95% CI:  $\pm 0$  year). Counter-factual given by a dynamic counterfactual forecast conditional on the control variables in absence of the detected break is shown as dashed lines for 5 years following the break. Introduction of the BC carbon tax is shown as gray-dotted line but is not detected as a break for BC in the model. Full estimation results are reported in Appendix I

### 3.3.3 Break Detection Estimates for CO<sub>2</sub> Emissions

Break detection<sup>14</sup> identifies three periods of significant province-level reductions in aggregate CO<sub>2</sub> emissions (Table 5 and Fig. 9), however, these coincide neither with the BC carbon tax, nor the introduction of carbon prices and trading in Alberta or Quebec. A significant decline in emissions is instead detected in New Brunswick in 2012 (95% CI:  $\pm 1$  year), Nova Scotia in 2012 (95% CI:  $\pm 3$  year) as well as in Newfoundland and Labrador in 1999 (95% CI:  $\pm 0$  year). The detected step-shift breaks are robust to allowing for the relationship

<sup>14</sup> Implemented using the gets package in R (Pretis et al. 2018a, b) setting  $p=0.025$  together with impulse indicator saturation (IIS: Hendry et al. 2008) as a robust estimator to remove potential outlying observations. See also Pretis et al. (2018a, b) as well as Nymoen and Sparman (2015) for an application of IIS in a fixed effects panel model.

between emissions and control variables to also change over time—when allowing for possible breaks in the coefficients on control variables, the identical treatment events are detected and no breaks on random variables are detected—see Appendix I). Attributing the detected step-shift breaks ex-post to possible ‘treatment interventions’, the decline in emissions in New Brunswick around 2012 is likely driven by reduced use of fossil-fuel driven power generators (see CBC 2012). In Nova Scotia, the detected break can be attributed to the closure of emission-intense pulp and paper mills together with their respective power plants<sup>15</sup> in 2012 (Govt. of Canada, 2014), as well as with electricity policies put into place in 2015 (Prov. Of Nova Scotia, 2017) which falls within the  $\pm 3$  year approximate 95% confidence interval on the detected break date (Hendry and Pretis 2016). The reduction in emissions in Newfoundland and Labrador in 1999 relative to control provinces coincides with increases in hydroelectric capacity for electricity generation (Newfoundland Power Emission Reduction Report, 2001). The detected breaks together with their predicted counter-factual over 5 years are shown in Fig. 9. The counterfactual is constructed as the dynamic forecast conditional on the observed control variables in absence of the detected break. Together these results show that industry-level interventions appear to have led to significant reductions in emissions while policy interventions in the form of carbon pricing have not had a detectable impact on aggregate emissions (Fig. 9).

Nevertheless, the introduction of the BC carbon tax is detected without prior knowledge of its occurrence in transportation emissions. Without *a-priori* specifying the BC carbon tax to be included in the model, break detection identifies the introduction of the tax as a significant reduction in BC transportation emissions in 2009 (95% CI:  $\pm 4$  year, see Appendix I). The detected break has an estimated magnitude of  $-5\%$  ( $se=2\%$ ) which is near identical to the difference-in-differences estimate imposing the intervention from the outset in transportation emissions.

## 4 Conclusion

Using difference-in-differences, synthetic control, and introducing a new econometric break detection approach to identify previously unknown treatment assignment and timing, I find evidence of some sectors having significantly reduced their CO<sub>2</sub> emissions in response to the introduction of a carbon tax in British Columbia, Canada. Particularly there is evidence for a significant reduction in emissions from transportation, the single largest source of emissions in the taxed region. However, there is no statistically significant effect of the introduction of the carbon tax on the *aggregate* level of CO<sub>2</sub> emissions, highlighting the heterogeneity in emission elasticities across sectors as well as the importance of command-and-control policy interventions beyond carbon pricing. These results are consistent across all methods employed.

The proposed panel econometric break detection approach identifies the BC carbon tax as having reduced transportation emissions. Detected significant reductions in *aggregate* CO<sub>2</sub> emissions across Canadian provinces, in turn, do not coincide with the implementation of the BC carbon tax, nor the carbon pricing schemes in Alberta and Quebec. Instead, the detected interventions in aggregate emissions reflect the closures of emission-intense industries (pulp and paper with accompanying power-generating facilities) and reforms in

<sup>15</sup> These are the Bowater Mersey and Minas Basin mills.

electricity generation in untaxed provinces such as New Brunswick and Nova Scotia transitioning away from fossil-fuel driven electricity generation. More generally, the break detection approach based on machine learning provides a new way to identify successful policy in settings where multiple (known and unknown) interventions have taken place.

How representative are the results from British Columbia for carbon taxes as a whole? The BC carbon tax enjoys a high degree of public acceptance and—with its relatively few exemptions—is a good candidate for the analysis of the tax impact on *aggregate* emissions. The effectiveness of reducing emissions in the transportation sector are encouraging, particularly as BC relies heavily on individual motor vehicles due to the long driving distances and limited public transport. This bodes well for the introduction of carbon taxes in comparable regions (such as the US). The significant impact of the tax on transportation emissions is also consistent with the impact of the Swedish carbon tax (Andersson 2019). BC, however, is a special case with regards to the tax impacts of electricity generation—the province relies predominantly on hydro power for electricity generation; thus it is not surprising to see little impact of the carbon tax on emissions related to electricity production.

A likely explanation for the absence of a detectable statistically significant tax effect on *aggregate* emissions is that the carbon tax is too low to result in rapid changes in the short time frame since its introduction given the heterogeneous emission elasticities across sectors. The break detection results also highlight that carbon taxes are not the only available tool to curb emissions—command and control approaches, closures of industries, and (policy mandated) efficiency improvements can lead to significant aggregate emission reductions. These findings also support the work of Eskander and Fankhauser (2020) arguing that climate policies have been limited in their large-scale effectiveness. Pricing carbon at currently-observed rates is unlikely to be sufficient alone to achieve major emission reductions and policy mixes will be needed.

Overall, the present results also highlight the importance of monitoring the effectiveness of policy tools based on their target metric—small changes in gasoline demand, transportation emissions, or natural gas use are insufficient when the required policy target is *aggregate* CO<sub>2</sub> emissions due to the cumulative nature of atmospheric CO<sub>2</sub> concentrations.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10640-022-00679-w>.

**Funding** Funding was supported by Robertson Foundation.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study of the Basque Country. *Am Econ Rev* 93(1):113–132
- Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J Am Stat Assoc* 105(490):493–505

- Abadie A, Diamond A, Hainmueller J (2011) Synth: an R-package for synthetic control methods in comparative case studies. *J Stat Softw* 42(13):10. <https://doi.org/10.18637/jss.v042.i13>
- Ahmadi Y (2017) How effective are carbon taxes in reducing emissions? Evidence from the Revenue Neutral Carbon Tax in British Columbia, Canada. Working Paper
- Ahmadi Y, Yamazaki A (2020) The Effectiveness of Revenue-neutral Carbon Tax (No. 19-36). National Graduate Institute for Policy Studies
- Akerlof GA (2020) Sins of omission and the practice of economics. *J Econ Lit* 58(2):405–418
- Andersson JJ (2019) Carbon taxes and CO<sub>2</sub> emissions: Sweden as a case study. *Am Econ J*
- Antweiler W, Gulati S (2016) Frugal cars or frugal drivers? How carbon and fuel taxes influence the choice and use of cars. Working Paper
- Apergis N, Lau MCK (2015) Structural breaks and electricity prices: further evidence on the role of climate policy uncertainties in the Australian electricity market. *Energy Econ* 52:176–182
- Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev Econ Stud* 58(2):277–297
- Azevedo D, Wolff H, Yamazaki A (2018) Do carbon taxes kill jobs? Firm-level evidence from British Columbia. Working Paper
- Best R, Burke PJ, Jotzo F (2020) Carbon pricing efficacy: cross-country evidence. *Environ Resource Econ* 77(1):69–94
- British Columbia (2008) Climate Action Plan. [https://www2.gov.bc.ca/assets/gov/environment/climate-change/action/cap/climateaction\\_plan\\_web.pdf](https://www2.gov.bc.ca/assets/gov/environment/climate-change/action/cap/climateaction_plan_web.pdf)
- Buchmueller TC, DiNardo J, Valletta RG (2011) The effect of an employer health insurance mandate on health insurance coverage and the demand for labor: evidence from Hawaii. *Am Econ J Econ Pol* 3(4):25–51
- Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. *Rev Econ Stat* 90(3):414–427
- Carattini S, Kallbekken S, Orlov A (2019) How to win public support for a global carbon tax. *Nature* 565:289–291
- Carbone J, Rivers N, Yamazaki A, Yonezawa H (2019) Comparing applied general equilibrium and econometric estimates of the effect of an environmental policy shock. *J Assoc Environ Resour Econ* 7:687
- Castle JL, Doornik JA, Hendry DF, Pretis F (2015) Detecting location shifts during model selection by step-indicator saturation. *Econometrics* 3(2):240–264
- CBC, Canadian Broadcasting Corporation (2012) N.B. leads pack for greenhouse gas emission reductions. <https://www.cbc.ca/news/canada/new-brunswick/n-b-leads-pack-for-greenhouse-gas-emission-reductions-1.1247654>
- Conley TG, Taber CR (2011) Inference with “difference in differences” with a small number of policy changes. *Rev Econ Stat* 93(1):113–125
- Croissant Y, Millo G (2008) Panel data econometrics in R: the plm package. *J Stat Softw.* <https://doi.org/10.18637/jss.v027.i02>
- Daniel KD, Litterman RB, Wagner G (2019) Declining CO<sub>2</sub> price paths. *Proc Natl Acad Sci* 116(42):20886–20891
- Davis LW, Kilian L (2011) Estimating the effect of a gasoline tax on carbon emissions. *J Appl Economet* 26(7):1187–1214
- Donaldson D, Storeygard A (2016) The view from above: applications of satellite data in economics. *J Econ Perspect* 30(4):171–198
- Elgie S, McClay J (2013) Policy Commentary/Commentaire BC’s carbon tax shift is working well after four years (attention Ottawa). *Can Public Policy* 39(Supplement 2):S1–S10
- Environment and Climate Change Canada (2018) Canada’s greenhouse gas inventory. <https://www.canada.ca/en/environment-climate-change/services/climate-change/greenhouse-gas-emissions/inventory.html>
- Ericsson NR (2017) How biased are US government forecasts of the federal debt? *Int J Forecast* 33(2):543–559
- Esarey J (2016) clusterSEs: calculate cluster-robust p-values and confidence intervals R package version 2.3.2
- Eskander SM, Fankhauser S (2020) Reduction in greenhouse gas emissions from national climate legislation. *Nat Clim Chang* 10(8):750–756
- Estrada F, Perron P, Martínez-López B (2013) Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nat Geosci* 6(12):1050
- Farmer JD, Hepburn C, Ives MC, Hale T, Wetzer T, Mealy P, et al. (2019) Sensitive intervention points in the post-carbon transition. *Science* 364(6436):132–134

- Feir D, Gillezeau R, Jones M (2018) Illuminating indigenous economic development (no. 1806). Department of Economics, University of Victoria
- Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 33(1):1
- Gaure S (2013) lfe: linear group fixed effects. *R Journal* 5(2):104–117
- Gittens J (2018) The impact of carbon taxes on market structure: evidence from British Columbia's Retail Gasoline Industry. Working Paper
- Government of Canada (2014) Regulatory impact analysis statement on coal-fired powerplant emissions in Nova Scotia. <http://www.gazette.gc.ca/rp-pr/p1/2014/2014-06-28/html/reg3-eng.html>
- Green JF (2021) Does carbon pricing reduce emissions? A review of ex-post analyses. *Environ Res Lett*
- Gulati S, Gholami Z (2015) Estimating the impact of carbon tax on natural gas demand in British Columbia. *Sustain Prosper*
- Hanoteau J, Talbot D (2019) Impacts of the Québec carbon emissions trading scheme on plant-level performance and employment. *Carbon Manag* 10(3):287–298
- Hansen BE (2017) Regression kink with an unknown threshold. *J Bus Econ Stat* 35(2):228–240
- Harrison K, Peet C (2012) Historical legacies and policy reform: diverse regional reactions to British Columbia's carbon tax. *BC Stud: Br Columbian Q* (173):97–122
- Hendry DF (2018) First-in, first-out: modelling the UK's CO<sub>2</sub> emissions, 1860–2016. Oxford Department of Economics Working Paper
- Hendry DF, Johansen S (2015) Model discovery and Trygve Haavelmo's legacy. *Econom Theory* 31(1):93–114
- Hendry DF, Pretis F (2016) Quantifying the uncertainty around break dates in step-indicator saturation. Oxford Department of Economics Working Paper
- Hendry DF, Santos C (2005) Regression models with data-based indicator variables. *Oxford Bull Econ Stat* 67(5):571–595
- Hendry DF, Johansen S, Santos C (2008) Automatic selection of indicators in a fully saturated regression. *Comput Stat* 23(2):317–335
- Janssens-Maenhout G, Pagliari V, Guizzardi D, Muntean M (2013) Global emission inventories in the emission database for global atmospheric research (EDGAR)—Manual (I). Gridding: EDGAR emissions distribution on global gridsmaps, Publications Office of the European Union, Luxembourg
- Johansen S, Nielsen B (2009) An analysis of the indicator saturation estimator as a robust regression estimator. In: Castle J, Shephard N (eds) *The methodology and practise of econometrics*, pp 1–36
- Johansen S, Nielsen B (2016) Asymptotic theory of outlier detection algorithms for linear time series regression models. *Scand J Stat* 43(2):321–348
- Koch N, Naumann L, Pretis F, Ritter N, Schwarz M (2022) What reduces road CO<sub>2</sub> emissions? Policy attribution using break detection. Working Paper
- Lawley C, Thivierge V (2016) Refining the evidence: British Columbia's carbon tax and household gasoline consumption
- Lin B, Li X (2011) The effect of carbon tax on per capita CO<sub>2</sub> emissions. *Energy Policy* 39(9):5137–5146
- MacKinnon JG, Webb M (2019) Randomization inference for difference-in-differences with few treated clusters (No. 1355). Queen's Economics Department Working Paper
- Metcalf GE (2019) On the economics of a carbon tax for the United States. *Brookings Papers on Economic Activity*, BPEA Conference Draft
- Metcalf GE, Stock JH (2020) The macroeconomic impact of Europe's carbon taxes. NBER working paper (w27488)
- Mukanjari S, Sterner T (2018) Do markets trump politics? Evidence from fossil market reactions to the Paris agreement and the US election
- Murray B, Rivers N (2015) British Columbia's revenue-neutral carbon tax: a review of the latest "grand experiment" in environmental policy. *Energy Policy* 86:674–683
- Newfoundland Power Emission Reduction Credit Creation Report (2001) [https://www.csaregistrries.ca/files/projects/prj\\_6300\\_668.pdf](https://www.csaregistrries.ca/files/projects/prj_6300_668.pdf)
- Nielsen B, Qian M (2018) Asymptotic properties of the gauge of step-indicator saturation. Working Paper
- Nickell S (1981) Biases in dynamic models with fixed effects. *Econometrica* 49:1417–1426
- Nymoen R, Sparman V (2015) Equilibrium unemployment dynamics in a panel of OECD countries. *Oxford Bull Econ Stat* 77(2):164–190
- Perron P (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57:1361–1401
- Piehl AM, Cooper SJ, Braga AA, Kennedy DM (2003) Testing for structural breaks in the evaluation of programs. *Rev Econ Stat* 85(3):550–558

- Pretis F, Schneider L, Smerdon JE, Hendry DF (2016) Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *J Econ Surv* 30(3):403–429
- Porter J, Yu P (2015) Regression discontinuity designs with unknown discontinuity points: testing and estimation. *J Economet* 189(1):132–147
- Province of Nova Scotia (2017) Environmental goals and sustainable prosperity act 2015–17 Progress Report. <https://novascotia.ca/nse/egspa/docs/EGSPA-2015-17-Progress-Report.pdf>
- Pretis F (2020) Econometric modelling of climate systems: The equivalence of energy balance models and cointegrated vector autoregressions. *J Economet* 214(1):256–273
- Pretis F, Schwarz M (2022) Discovering what mattered: answering reverse causal questions by detecting unknown treatment assignment and timing as breaks in panel models. Working Paper
- Pretis F, Mann ML, Kaufmann RK (2015) Testing competing models of the temperature hiatus: assessing the effects of conditioning variables and temporal uncertainties through sample-wide break detection. *Clim Change* 131(4):705–718
- Pretis F, Reade JJ, Sucarrat G (2018a) Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *J Stat Softw.* <https://doi.org/10.18637/jss.v086.i03>
- Pretis F, Schwarz M, Tang K, Hausteine K, Allen MR (2018b) Uncertain impacts on economic growth when stabilizing global temperatures at 1.5° C or 2° C warming. *Philos Trans R Soc A* 376:20160460
- Rafaty R, Dolphin G, Pretis F (2020) Carbon pricing and the elasticity of CO<sub>2</sub> emissions. INET Economics Working Paper
- Rivers N, Schaufele B (2015) Saliency of carbon taxes in the gasoline market. *J Environ Econ Manag* 74:23–36
- Ruhm CJ (2019) Shackling the identification police? *South Econ J* 85(4):1016–1026
- Saberian S (2018) The negative effect of the BC carbon tax on vancouver air quality: a good climate for bad air? Working Paper
- Schneider L, Smerdon JE, Pretis F, Hartl-Meier C, Esper J (2017) A new archive of large volcanic events over the past millennium derived from reconstructed summer temperatures. *Environ Res Lett* 12(9):094005
- Simes RJ (1986) An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73(3):751–754
- Statistics Canada (2018) Province or territory data. <https://www150.statcan.gc.ca/n1/en/type/data?MM=1>
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc B* 58(1):267–288
- Walker AJ, Pretis F, Powell-Smith A, Goldacre B (2019) Variation in responsiveness to warranted behaviour change among NHS clinicians: novel implementation of change detection methods in longitudinal prescribing data. *bmj*, 367
- Williams LV, Reade JJ (2016) Prediction markets, social media and information efficiency. *Kyklos* 69(3):518–556
- Xiang D, Lawley C (2018) The impact of British Columbia's carbon tax on residential natural gas consumption. *Energy Economics*
- Yamazaki A (2017) Jobs and climate policy: evidence from British Columbia's revenue-neutral carbon tax. *J Environ Econ Manag* 83:197–216
- Yip CM (2018) On the labor market consequences of environmental taxes. *J Environ Econ Manag* 89:136–152
- Zhang Q, Pandey B, Seto KC (2016) A robust method to generate a consistent time series from DMSP/OLS nighttime light data. *IEEE Trans Geosci Remote Sens* 54(10):5821–5831

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.