Netherlands Organisation for Scientific Research
Social Sciences

Young Researcher Causal Inference
Application Form 2023/2024

Registration form

1a. Details of applicant
-Name, title(s): Khanh Nguyen, Takeshi Takiguchi
-University, Department (or Institute): Vrije Universiteit Amsterdam
-E-mail(s): duykhanh.nguyen@student.kuleuven.be, takeshi.takiguchi@wur.nl

1b. Title of research proposal
Synthetic Control Method for Carbon Tax Effectiveness Evaluation in British Columbia

1c. Abstract
This research proposal examines the effect of carbon tax on CO2 emissions in British Columbia, utilizing the synthetic control method. While previous studies using this method have questioned the tax's effectiveness, there are some limitations on their predictor selection and inference methods. This proposed study aims to refine these methodologies by applying cross-validation for predictor selection, a t-test on k-fold cross-validation to determine statistical significance, and generalized synthetic control methods to compare treatment effects. Provisional results indicate a statistically significant increase in CO2 emissions in British Columbia relative to the synthetic control after the carbon tax was implemented. However, the generalized synthetic control methods reveal potential violations of the parallel trend assumption, suggesting further investigation is needed to establish a reliable counterfactual for British Columbia. Future analyses will focus on constructing synthetic controls on longer pre-tax periods, exploring additional predictors, and implementing 2-stage optimization of region and predictor weights.

Word Count: 3719
OSF project DOI: 10.17605/OSF.IO/247UE
Github URL: https://github.com/khanhnguyendata/vuass_carbon_tax_synthetic_control/

Research proposal

2a. Description of the proposed research and societal results

3000-4000 words for sections 2a-c (excluding references, tables, and figures, but including footnotes) (use Word Count to specify number of words)

2.a.i Introduction

To achieve the global goal of limiting temperature rise to less than 2.0 degrees Celsius while pursuing efforts to 1.5 degrees as set in the Paris Agreement, the world needs to accelerate its efforts to reduce greenhouse gas (GHG) emissions. While there exist various policy tools to cut GHG emissions, carbon pricing is considered an effective policy instrument to reduce emissions over the years. The major two instruments of carbon pricing are carbon tax and emissions trading system (ETS). Carbon tax has been implemented by 39 countries including Chile, New Zealand, Norway and South Africa, and an increasing number of countries are considering or under development of carbon tax as well (World Bank, 2024). Similarly, ETS is expanding worldwide, with EU-ETS recognized as the most influential and effective example of ETS ever implemented.

Canada is one of the top GHG emitters in the world, producing approximately 750 million tons of CO2 equivalent in 2022 (European Commission et al., 2023). Among 10 provinces and 3 territories within the country, British Columbia is considered as the forefront and the most ambitious province to tackle climate change. The province implemented a comprehensive carbon tax in 2008 and it was the first province or state that implemented a widespread carbon tax in North America (Pretis, 2022). The tax is placed on purchase and use of fossil fuels and covers nearly 80% of the province's emissions (Government of British Columbia, 2024). It has gradually increased from 10 Canadian Dollars per ton of CO2 emitted, to 80 Canadian Dollars in 2024 (Arcila & Baker, 2022; Government of British Columbia, 2024). Revenues obtained from carbon tax are used to fund environmental initiatives and mitigate the negative impact caused by the tax on low- and middle-income households (Arcila & Baker, 2022).

2.a.ii Previous Studies

Due to its pioneer role in tackling climate change using carbon tax, British Columbia has gained considerable attention from a number of researchers and its effect on reducing GHG emissions has been widely examined. An extensive amount of research has shown positive impacts of carbon tax on CO2 emissions reduction and the carbon tax in British Columbia has been considered as a successful story to tackle climate change (Murray & Rivers, 2015). However, recent literature has shown that a carbon tax has little or even negative impact on reducing CO2 emissions using synthetic control methods (Arcila & Baker, 2022; Pretis, 2022). These studies pointed out that British Columbia's carbon tax rate is too low to be effective in reducing CO2 emissions, amongst other reasons.

First, Arcila & Baker (2022) used pre-tax CO2 emissions, log population, real GDP growth, unemployment rate, and the share of employment in the fossil fuel sector from 60 other states & provinces in Canada and the United States to construct the synthetic control for British Columbia. They showed that CO2 emissions have actually increased after the implementation of carbon tax in British Columbia relative to the synthetic control.

More specifically, a difference-in-difference comparison between British Columbia and the synthetic control shows that while British Columbia experienced on average a 2.1 decrease in emission index from pre-tax (1998-2007) to post-tax (2008-2017) period, with emission index taken as the annual CO2 emission in British Columbia relative to province's CO2 emission in 1998 (which takes a base value of 100). In contrast, the synthetic control experienced on average a 6.1 decrease in emission index across the same periods. This implies an average 4.0 increase in emission index experienced by British Columbia compared to the synthetic control after the carbon tax was rolled out in 2008, an amount equivalent to 4.0% of the CO2 emissions by British Columbia in 1998.

Second, Pretis (2022) used CO2 emissions for the outcome variable, and GDP and population data as predictors to construct the synthetic control, covering the time frame from 1990 to 2016. While he found that there was a significant emissions reduction in the transport sector, he shows that the tax has not had a statistically significant impact on overall CO2 emissions reduction in British Columbia compared to the synthetic control.

While those studies are noteworthy for casting doubt on the effectiveness of the carbon tax in British Columbia for CO2 emissions reduction, their analyses could be further refined. First, it is not clear how the predictors for the synthetic control were chosen. For instance, Arcila & Baker (2022) chose population, GDP growth rate, unemployment rate, and share of employment in the energy industry as predictors, but the selection is not clearly articulated in the paper. They also used CO2 emissions data from 2000, 2004 and 2008 as lagged predictors, but these time periods seem to have been chosen without any explanation.

Second, these findings lack informative statistical significance of the effect of carbon tax on CO2 emission. For instance, Pretis (2022) used a permutation test to determine the statistical significance of the effect of carbon tax on CO2 emission (Abadie et al., 2010). However, with a small sample size, the permutation test cannot serve as an appropriate way to make inference (Hahn & Shi, 2017): as his control group only includes 9 Canadian provinces other than British Columbia, he acknowledges that the p-value obtained from the test (with a value of 1) is only marginally useful. In contrast, Arcila & Baker (2022) uses bootstrapped standard errors from 399 repetitions to confirm the increase in CO2 emission of British Columbia relative to the synthetic control is statistically significant. However, they did not specify in detail how these bootstrapped standard errors were obtained nor provided any reference to their inference method.

By building upon those studies, the aim of our research is to explore the effect of carbon tax in British Columbia on reduction in CO2 emissions, using the synthetic control method with some improved methodologies.

2b. Approach: what is your methodological or experimental approach?
For this project, we plan to reuse the data & analysis from Arcila & Baker in their 2022 study. As a result, the following predictors for each region (state or province) were used in their study to construct the synthetic control for British Columbia in the period before the carbon tax was implemented in 2008:

- Emission index in 2000, 2004, and 2008 (lagged outcomes)
- Average real GDP growth from 1998 to 2008
- Average unemployment rate from 1998 to 2008
- Log of population in 2008
- Share of employment in the energy industry in 2008

Using these predictors, Arcila & Baker (2022) constructed a synthetic control for British Columbia from the 60 other states and provinces in the United States & Canada, such that it is as close as possible to British Columbia in terms of these predictors. This means minimizing a weighted average of least square differences between British Columbia and the synthetic control across the chosen predictors, with the predictor weights constrained between 0 and 1 and adding up to 1. In turn, the predictor value for the synthetic control is a weighted average of the predictor value across the regions, with the region weights constrained between 0 and 1 and adding up to 1.

The optimal predictor weights were chosen such that the optimal region weights obtained by minimizing the least square difference in predictors also minimizes the least square difference in the pre-tax outcome (CO2 emission) between British Columbia and the synthetic control. This implies that the synthetic control, being similar to British Columbia in the relevant predictors, is also similar to British Columbia in the pre-tax outcomes. Hence, it can serve as a reliable counterfactual for British Columbia post-tax.

In our study, we plan to apply the following methodologies on top of the results from Arcila & Baker to address the aforementioned limitations of their & similar studies:

A. Cross-validation to determine which predictors to include in the synthetic control by their predictive power of the pre-tax CO2 emissions in British Columbia (Abadie, 2021).
B. t-test on k-fold cross-validation to determine statistical significance of the effect of carbon tax on CO2 emission in British Columbia relative to the synthetic control (Chernozhukov, 2018)
C. Generalized synthetic control methods such as fixed effect and interaction fixed effect model to compare with the result from the t-test on our synthetic control (Xu, 2017)

2c. Provisional results: What are your findings?
A. Cross-validation for predictor selection of synthetic control
Before we select different predictors for the synthetic control, we decided to restrict the predictors to end before the tax was implemented in 2008. In other words, predictors that involve the year 2008 were replaced by the year 2007, such as emission index in 2007, or the average GDP growth from 1998 to 2007. In contrast, Arcila & Baker (2022) uses predictors that end in 2008, which might result in a biased synthetic control: for instance, the synthetic control was constructed to be similar to British Columbia in 2008 CO2 emission index (one of the predictors used to construct the synthetic control), even though it should be different from British Columbia given the province rolled out the carbon tax in the middle of 2008.

When the last year of the pre-tax period for the predictors was changed from 2008 to 2007, the effect of the carbon tax on CO2 emission of British Columbia increased from 4.56 to 6.23. Note that the former treatment effect of 4.56 is slightly different from the 4.0 treatment effect from Arcila & Baker (2022). This could be due to the different synthetic control algorithms implemented in Stata (used by Arcila & Baker) versus R (used by us, with care to ensure our R code is as similar as possible to the authors' Stata code), a discrepancy that was also documented by Becker & Klößner (2018). This could also be due to the non-unique weights from traditional synthetic control algorithms, which can vary wildly even from innocuous re-orderings of the units (Klößner et al., 2018): we also found that the weights–and to a lesser extend the treatment effect–can vary wildly even when we simply imputed some missing data.

With the modified predictors that ended in 2007, we experimented with different combinations of predictors and evaluated their predictive power on the treatment outcome using 2-fold cross-validation. More specifically (see Figure 1):

- We divide the pre-tax period into 2 smaller periods: a "training period" from 1998 to 2003, and a "validation period" from 2004 to 2007
- We construct the synthetic control using predictors from the training period and calculate the root mean square prediction error between the CO_2 emission of British Columbia and the synthetic control during this period (training RMSPE).
- We use the synthetic control to predict the CO_2 emission index of British Columbia in the validation period, and calculate the root mean square prediction error between the predicted and actual value during this period (validation RMSPE). Since the validation period is still before the carbon tax came into effect, a reliable synthetic control should a priori have a low validation RMSPE i.e. its CO_2 emission index should be close to that of British Columbia during this period.

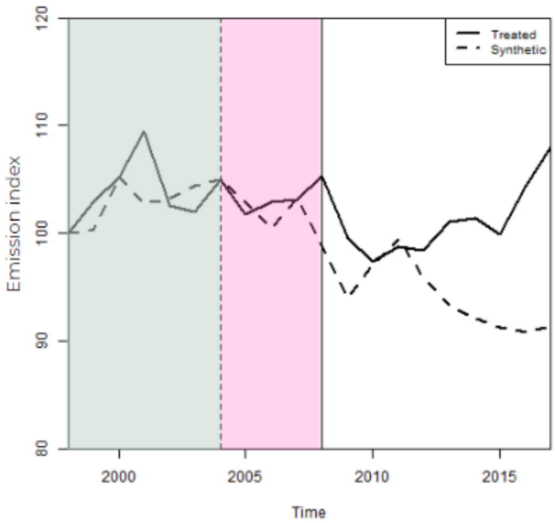

Figure 1: 2-fold cross-validation scheme for predictor selection, highlighting training period (green) & validation period (pink)

The combinations of predictors are chosen across 2 dimensions:

- Outcome type: type of lagged outcome predictors (emission index), adapted from Kaul et al. (2022)
    - *all*: all lagged outcomes (1999*-2003)
    - *first_mid_list*: lagged outcomes near the first, middle, and last years of the training period (1999*, 2001, 2003). This is similar to the configuration used by Arcila & Baker (2022).
    - *last*: only the lagged outcome at the last year of the training period (2003)
    - *mean*: average outcome across the training period (1998-2003)
- Additional predictor type: type of predictors in addition to the lagged outcomes
    - *none*: no additional predictor
    - *econ*: economic & demographic predictors used in the original study (average real GDP growth from 1998 to 2008, average unemployment rate from 1998 to 2008, log of population in 2008)
    - *all*: all additional predictors in the original study, including the economic & demographic predictors above, as well as energy-related predictor (share of employment in the energy industry in 2008)

* The first lagged outcome of 1998 is not included as predictor since it has no variance: all regions have emission index of 100 in 1998

Comparing the training & validation RMSPE across these predictor combinations (Table 1), we notice that:

- Having only the last lagged outcome & no additional predictor has the lowest validation RMSPE (1.60)
- When more lagged outcomes are added to this configuration (such as having all lagged outcomes), the training RMSPE decreases (0.01 vs 3.75), while the validation RMSPE increases (3.50 vs 1.60). This implies that having more lagged outcomes as predictors tend to overfit the synthetic control, leading to its worse predictive power.
- Similarly, when more predictors are added to this configuration (such as having all additional predictors), the training RMSPE decreases (2.55 vs 3.75), while the validation RMSPE increases (2.69 vs 1.60). This also implies that having more additional perdictors tend to overfit the synthetic control, leading to its worse predictive power.

Training RMSPE	*Additional predictor type*		
Outcome type	none	econ	all
all	0.01	0.02	0.02
first_mid_last	0.47	0.08	0.53
last	3.75	2.67	2.55
mean	3.24	2.64	2.08
Validation RMSPE	*Additional predictor type*		
Outcome type	none	econ	all
all	3.50	3.49	3.48
first_mid_last	3.26	3.74	1.84
last	1.60	1.94	2.69
mean	3.98	3.33	4.41

Table 1: Training & validation RMSPE for different outcome types & additional predictor types

Therefore, the predictor that we choose for our synthetic control will only be the last emission index in 2007. It is also consistent with the recommendation from Kaul et al. (2022) of preferring the 'last' lagged outcome configuration when there are also other relevant predictors, even though in this case it also has good predictive power without any additional predictor.

B. <u>T-test on k-fold cross-validation to determine statistical significance of the effect of carbon tax</u>
With the synthetic control constructed using only the last emission index in 2007, we obtained a treatment effect of 6.28, compared to 6.23 using the original predictors ending in 2007. This large positive treatment effect can also be seen in Figure 2, where the post-tax emission index of British Columbia is much higher than the synthetic control, hence a large positive difference in post-tax emission index.

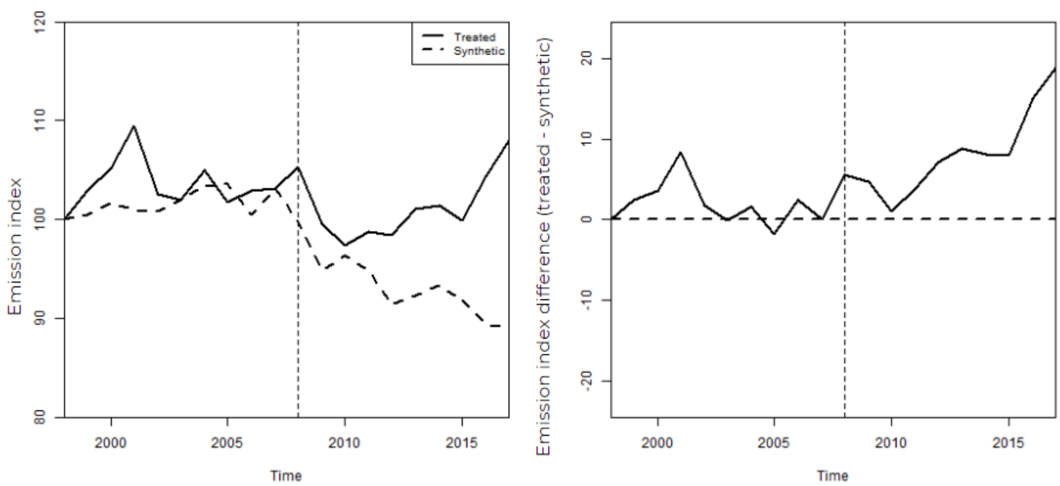

Figure 2: Left: Emission index of British Columbia (Treated) and the synthetic control (Synthetic). Right: Difference in emission control of British Columbia compared to the synthetic control

To determine the statistical significance of this treatment effect, we follow the k-fold cross-validation strategy proposed by Chernozhukov et al. (2018), where the number of fold (k) is chosen as 3 from the authors' recommendation for small number of pre-treatment periods. More specifically (Figure 3):

- We divide the pre-treatment period into 3 smaller periods (folds): 1998-2000, 2001-2003, and 2004-2007
- For each fold, we construct a synthetic control using the last emission index in the fold as predictor, and evaluate the treatment effect of the carbon tax on British Columbia relative to the synthetic control in the post-tax period.
- We average the treatment effects across folds to get an average treatment effect of carbon tax. In addition, we calculate the variance across these treatment effects to determine the standard deviation, hence p-value, of the treatment effect.

Figure 3: A synthetic control is constructed from each fold (green) and its corresponding treatment effect is measured in the post-tax period (pink)

Under 3-fold cross-validation, the average treatment effect of carbon tax on British Columbia relative to the synthetic control is 5.69. While this is slightly different from the treatment effect of 6.28 for the synthetic control constructed on the full pre-tax period (1998-2007), it allows us to obtain the p-value of 0.037 for this treatment effect. In other words, British Columbia has a statistically significant increase in CO2 emission index relative to synthetic control after the carbon tax was rolled out.

C. <u>Generalized synthetic control methods to compare with the result from the t-test</u>

While the treatment effect from the t-test is consistent in terms of statistical significance to the original study from Arcila & Baker (2022), albeit with a higher point estimate, we'd like to compare its result to alternative synthetic control methods, such as the more recent generalized synthetic control methods (Xu, 2017).

To this end, we used 2 different generalized synthetic control methods: the fixed effects (FE) and the interactive fixed effects (IFE) models. Comparing their result to the t-test result (Table 2), we see that the treatment effect from the t-test (5.69) is quite similar to that from the FE & IFE models (5.45). Furthermore, all three treatment effects are statistically significant, even after applying Holm's correction for multiple comparisons (Holm, 1979). This lends more credence to the statistically significant positive effect of carbon tax on the CO2 emission index of British Columbia under our synthetic control method.

Method	ATT	Lower CI (95%)	Upper CI (95%)	p-value
FE	5.45	2.58	8.32	0.00
IFE	5.45	2.91	7.99	0.00
T-test	5.69	0.60	10.78	0.04

Table 2: Comparison of treatment effect on emission index across 3 synthetic control methods, with corresponding lower & upper 95% confidence interval and p-value

On the other hand, the generalized synthetic control methods appear to violate the parallel trend assumption when testing them under placebo treatments (Figure 4). In other words, these synthetic control methods do not seem to provide reliable counterfactuals for British Columbia when no carbon tax was applied before 2008 (placebo treatment): the outcome predictions from the synthetic control is significantly different from the actual outcome of British Columbia during the pre-tax period.

Furthermore, these parallel trend violations appear to be larger than the lower bound of the treatment effect. This means the observed positive treatment effect could be largely influenced by these parallel trend violations, and might not reflect the true effect of carbon tax. Therefore, we need to exercise caution when drawing conclusions on the treatment effect from these synthetic control methods.

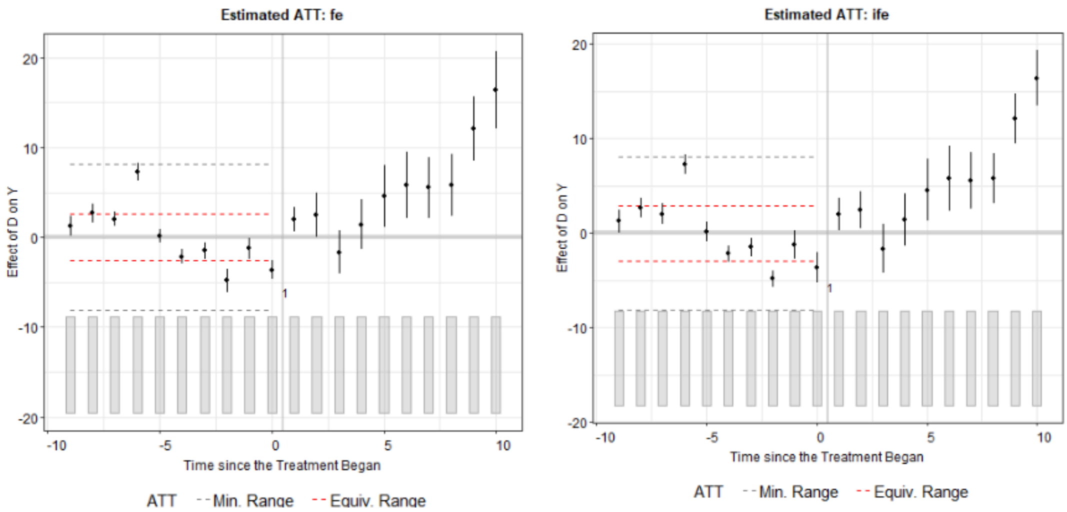

Figure 4: Degree of parallel trend assumption violation (gray dotted line) against the lower bound of the observed treatment effect (red dotted line) for generalized synthetic control methods using emission index

One potential explanation for the unreliability of the synthetic control methods is the potential effect of GDP on emission: high industrial output tends to correlate with high CO2 emission. This can be seen when comparing the real GDP of British Columbia against that of the synthetic control, indexed on the 1998 value for each unit. Real GDP of British Columbia tends to grow much more than the synthetic control, especially in the post-tax period (Figure 5). This could partly explain why, aside from the effect of the carbon tax, the post-tax CO2 emission of British Columbia is also much higher relative to the synthetic control, despite the two units having similar CO2 emission pre-tax (Figure 2).

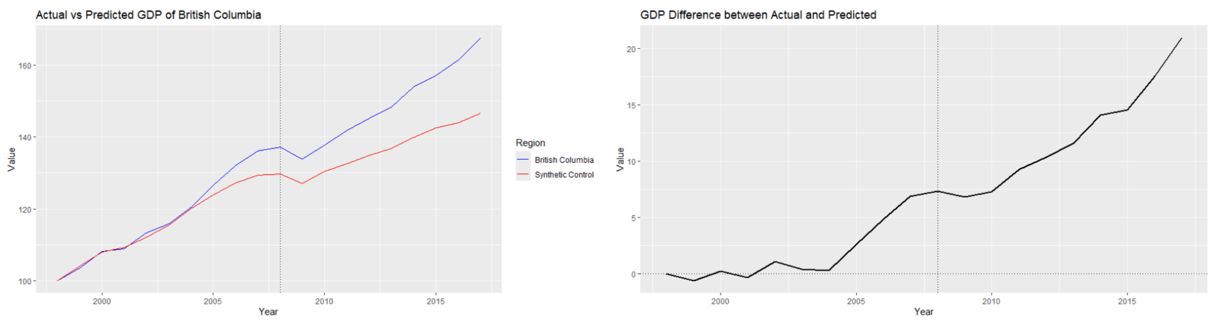

Figure 5: Real GDP of British Columbia vs synthetic control, indexed to 1998 values (left), and their differences (right)

A simple way to control for the effect of GDP on emission is to construct the synthetic control using emission per GDP rather than emission indexed to its initial value at 1998 (similar to per capita outcomes typically used in the synthetic control literature). The synthetic control using emission per GDP in 2007 as the sole predictor results in a negative treatment effect of carbon tax on CO2 emission per GDP of British Columbia compared to the synthetic control (Figure 6), which is not statistically significant under the t-test (Table 3). In contrast, the generalized synthetic control methods show a more negative effect of carbon tax on emission per GDP for British Columbia, which are also statistically significant.

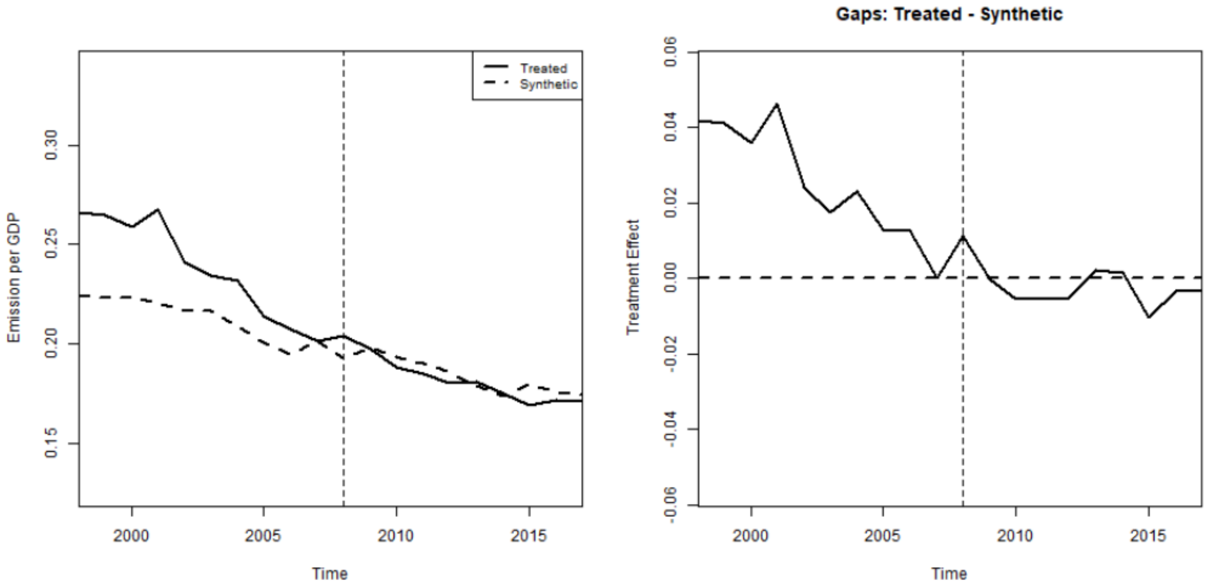

Figure 6: Left: Emission per GDP of British Columbia (Treated) and the synthetic control (Synthetic). Right: Difference in emission per GDP of British Columbia compared to the synthetic control

Method	ATT	Lower CI (95%)	Upper CI (95%)	p-value
FE	-0.045	-0.053	-0.037	0.000
IFE	-0.045	-0.053	-0.036	0.000
T-test	-0.026	-0.054	0.002	0.061

Table 3: Comparison of treatment effect on emission per GDP across 3 synthetic control methods, with corresponding lower & upper 95% confidence interval and p-value

While further investigation needs to be done to reconcile the result from the different synthetic control methods, the parallel trend assumption violation for the generalized synthetic control methods are less severe when using emission per GDP as outcome compared to using emission index (Figure 7). On the other hand, there are still significant violations that seem to trend downward, which could indicate that the synthetic control is still not a reliable counterfactual for British Columbia. It is also consistent with the poor fit of the synthetic control on the pre-tax emission per GDP of British Columbia, as seen in Figure 6.

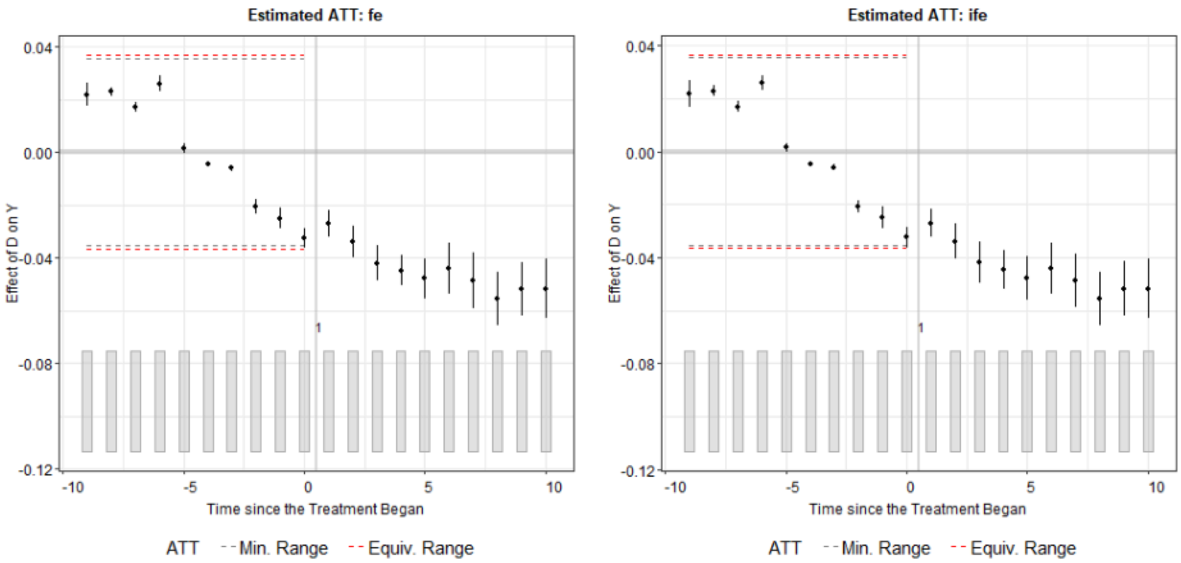

Figure 7: Degree of parallel trend assumption violation (gray dotted line) against the lower bound of the observed treatment effect (red dotted line) for generalized synthetic control methods using emission per GDP

This leads to the following analyses that we would like to explore next in our studies:

- Construct synthetic control on longer pre-tax periods, as typically recommended for synthetic control methods (Abadie, 2021). For instance, Pretis (2022) uses data from 1990 onward to construct their synthetic control for British Columbia.
- Explore & validate additional predictors, given that solely using the last emission per GDP in the pre-tax period might not provide a reliable counterfactual to British Columbia
- Implement 2-stage optimization of region & predictor weights to minimize both training and validation RMSPE (Becker et al., 2018), which could result in more reliable synthetic control in terms of predictive power of the outcome under cross-validation/placebo treatment
- Reconcile the different results on treatment effect of carbon tax on CO_2 emission of British Columbia using synthetic control between Arcila & Baker (2022) and Pretis (2022)

Word Count (Sections 2a – 2c): 3277

References

You must cite 5-10 academic sources with proper DOIs.

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. Journal of Economic Literature, 59(2), 391–425. https://doi.org/10.1257/jel.20191450

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association, 105(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Arcila, A., & Baker, J. D. (2022). Evaluating carbon tax policy: A methodological reassessment of a natural experiment. Energy Economics, 111, 106053. https://doi.org/10.1016/j.eneco.2022.106053

Becker, M., & Klößner, S. (2018). Fast and reliable computation of generalized synthetic controls. Econometrics and Statistics, 5, 1–19. https://doi.org/10.1016/j.ecosta.2017.08.002

Becker, M., Kloßner, S., & Pfeifer, G. (2018). Cross-Validating Synthetic Controls. https://mpra.ub.uni-muenchen.de/83679/1/MPRA_paper_83679.pdf

Chernozhukov, V., Wuthrich, K., & Zhu, Y. (2018). A t-test for synthetic controls (Version 8). arXiv. https://doi.org/10.48550/ARXIV.1812.10820

European Commission. Joint Research Centre. (2023). GHG emissions of all world countries: 2023. Publications Office. https://data.europa.eu/doi/10.2760/953322

Hahn, J., & Shi, R. (2017). Synthetic Control and Inference. Econometrics, 5(4), 52. https://doi.org/10.3390/econometrics5040052

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6(2), 65–70.

Kaul, A., Klößner, S., Pfeifer, G., & Schieler, M. (2022). Standard Synthetic Control Methods: The Case of Using All Preintervention Outcomes Together With Covariates. Journal of Business & Economic Statistics, 40(3), 1362–1376. https://doi.org/10.1080/07350015.2021.1930012

Klößner, S., Kaul, A., Pfeifer, G., & Schieler, M. (2018). Comparative politics and the synthetic control method revisited: A note on Abadie et al. (2015). Swiss Journal of Economics and Statistics, 154(1), 11. https://doi.org/10.1186/s41937-017-0004-9

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. https://doi.org/10.1016/j.enpol.2015.08.011

Pretis, F. (2022). Does a Carbon Tax Reduce CO2 Emissions? Evidence from British Columbia. Environmental and Resource Economics, 83(1), 115–144. https://doi.org/10.1007/s10640-022-00679-w

Strategy, M. of E. and C. C. (2024, April 2). British Columbia's Carbon Tax—Province of British Columbia. Province of British Columbia. https://www2.gov.bc.ca/gov/content/environment/climate-change/clean-economy/carbon-tax

World Bank. (2024). State and Trends of Carbon Pricing. https://openknowledge.worldbank.org/server/api/core/bitstreams/253e6cdd-9631-4db2-8cc5-1d013956de15/content

Xu, Y. (2017). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. Political Analysis, 25(1), 57–76. https://doi.org/10.1017/pan.2016.2