The Kyoto Protocol and Carbon Dioxide Emissions in New Zealand: A Synthetic Control Approach

Mitchell Pennington

A thesis submitted to Auckland University of Technology in partial fulfillment of the requirement for the degree of Master of Business (MBus)

2021

School of Economics

Faculty of Business, Economics and Law

Abstract

On the 11th of December 1997 the International Community signed the Kyoto Protocol: an international environmental treaty that commits countries to reducing their greenhouse gas (GHG) emissions to mitigate the effects of human activity driven climate change. The Kyoto Protocol imposes individual GHG emissions reductions targets on developed countries for commitment period one (2008-2012). Reduction targets amount to an aggregate 5% reduction in GHG emissions for participating countries when compared to 1990 levels of GHG emissions. The Kyoto Protocol is criticised as insufficient, with criticisms focusing on its structure. The inclusion of flexibility mechanisms, unrestricted international emissions trading and the large endowment of emissions credits given to former Soviet Union countries are said to have created compliance costs that fail to encourage any real decrease in emissions. The withdrawal of the United States from the Kyoto Protocol in 2001, the largest GHG emitter at the time, furthered the worries that the Kyoto Protocol would result in "business-as-usual" emissions. I analyse the effects of a legally binding emissions reductions target on the carbon dioxide (CO_2) emissions of New Zealand. Formally I seek to answer if the legally binding emissions reductions targets of the Kyoto Protocol reduce the carbon dioxide emissions of New Zealand? I employ the Synthetic Control Method (SCM), in which a weighted average of an untreated series is used as a counterfactual estimate to a treated series, to estimate the CO_2 emissions of New Zealand had it not joined the Kyoto Protocol. Furthermore, I extended the set of variables previously used to estimate causal effects of the Kyoto Protocol to include variables that are considered crucial in forecasting GHG emissions in Computable General Equilibrium (CGE) analyses. I then seek to answer a second question, are causal analysis specifications for the Kyoto protocol improved be including variables used to forecast emissions in CGE modelling? My results provide no statistically significant evidence that legally binding emissions targets result in New Zealand experiencing a reduction in CO_2 emissions. Additionally, I show that the inclusion of common variables in CGE literature, when estimating Synthetic Controls for climate policy, provides a better pre-treatment fit and bring results closer to statistical significance. Furthermore, my results show that, contrary to previous work, the use of US state level data is not always preferable to country level data when using the SCM to estimate causal effects of climate policy.

Table of Contents

Abstract 2

List of Figures 6

List of Tables 6

Attestation of Authorship 7

Acknowledgments 8

List of Acronyms 9

Section 1 Introduction 10

Section 2 Background 12

2.1 Climate Change 12

2.2 UNFCCC 13

2.3 Kyoto Agreement 14

2.4 New Zealand 16

Section 3 Literature Review 19

3.1 Theoretical Models 19

3.1.1 The income-environment relationship 21

3.1.2 Carbon Leakage 21

3.1.3 Overview of CGE/Theory 22

3.2 Causal Analysis Literature 23

3.3 Variable Choice in Causal Analysis and CGE literature 26

Section 4 Methodology 28

4.1 Causal Estimation 28

4.2 The Synthetic Control Method 30

4.2.1 Determining Weights 31

4.3 Significance Tests 33

4.4 Contextual Requirements 34

4.4.1 Size of effect and volatility of the outcome 34

4.4.2 Relevant Comparison Group Available 34

4.4.3 No Anticipation 35

4.4.4 No Interference 35

4.4.5 Convex Hull Condition 36

4.5 Robustness Checks 36

Section 5 Data 37

5.1 Outcome Variable 37

5.2 Country level data 38
5.2.1 Countries of Interest 39
5.2.2 Country level descriptive statistics 40
5.3 US state level data 41
5.3.1 US state level descriptive statistics 42
5.4 SCM vs Entire Donor Pool 43
5.5 Variable Descriptions 44
Section 6 Results 46
6.1 Primary Specification 46
6.2 Control for CDM's 48
6.3 Leave-One out tests 50
6.4 US state level data analysis 52
6.5 Results summary 53
Section 7 Discussion 54
Section 8 Conclusion 57
References 58
Appendices 62
Appendix A. Treatment effects by year 62
Appendix B. Recreating the Almer and Winkler (2017) result of Australia 63
Appendix C. Comparing the specifications using Australia as the treated country 64
Appendix D. Comparison of V weight methods 66

List of Figures

Figure 1. CO_2 emissions (kt) of Russia ... 15

Figure 2. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B donor pool) ... 46

Figure 3. Results of the Placebo tests (non-Annex B donor pool) ... 47

Figure 4. The p-value plots (non-Annex B donor pool) ... 47

Figure 5. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B CDM excluded donor pool) ... 48

Figure 6. Results of the Placebo tests (non-Annex B CDM excluded donor pool) ... 49

Figure 7. The p-value plots (non-Annex B CDM excluded donor pool) ... 50

Figure 8. Results of the "Leave-one out" Tests (non-Annex B CDM excluded donor pool) ... 51

Figure 9. The Synthetic and Actual CO_2 emissions for New Zealand (US state donor pool) ... 52

List of Tables

Table 1. Timeline of International Environmental Policy ... 18

Table 2. Summary of the Causal Analysis Literature ... 26

Table 3. Variables used to predict/model emissions ... 26

Table 4. Pairwise Correlation test ... 37

Table 5. Non-Annex B Donor Pool ... 39

Table 6. Country level descriptive statistics ... 40

Table 7. Descriptive Statistics Mean Comparisons ... 41

Table 8. US State level Descriptive Statistics ... 42

Table 9. Predictor Balance ... 43

Table 10. Descriptive statistics details ... 44

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly referenced), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Student's signature:

Date: 13 August 2021

Student ID: 1302215

Acknowledgments

I would like to express my gratitude to my supervisors, Peer Skov and Niven Winchester, who have provided me with invaluable support throughout completing this thesis. Thank you for providing constant encouragement and valuable input into this project, without it I would have been lost without a cause.

I would also like to thank my postgrad cohort. You all gave valuable insights into the material while also providing plenty of laughs along the way. We all worked incredibly well together, even during covid, and I sincerely hope we get to collaborate on work in our professional lives.

To my friends and family. Thank you for all your support throughout this journey. You all helped me in your own way, trying to make my life easier during this process in any way you could. It is more than I could ever ask for and I am forever grateful to have you all in my life.

Finally, to my partner KK. Thank you for always being there for me and helping me when it all felt overwhelming. You have gone beyond and above for me throughout my entire post grad experience. You were with me through my stress and angst, my happiness and joy, my setbacks and wins. I could not have asked for a better person to experience this journey with, and I will strive to support you in similar ways in the future.

Honourable mention to coffee.

List of Acronyms

AAU	assigned amount unit
CDM	clean development mechanism
CER	certified emissions reduction
CGE	computable general equilibrium
CO_2	carbon dioxide
COP	conference of the parties
CP1	commitment period 1
DiD	difference in differences
EIT	economy in transit
EKC	environmental Kuznets curve
ERU	emissions reduction unit
FSU	former Soviet Union
GDP	gross domestic product
GHG	greenhouse gas
HCI	human capital index
JI	joint implementation
NZ ETS	New Zealand emissions trading scheme
OECD	Organisation for Economic Co-operation and Development
PWT	Penn World Tables
RMU	removal unit
SCM	synthetic control method
UNFCCC	United Nations Framework Convention on Climate Change
US	United States of America
WDI	World Development Indicator

Section 1 Introduction

Climate changed caused by rising greenhouse gas (GHG) emissions is an issue faced worldwide (Field et al., 2014). Recognizing the need for collective action to combat climate change, 185 countries – including New Zealand – adopted 'The United Nations Framework Convention on Climate' in 1992. A subsidiary agreement, The Kyoto Protocol, was adopted on 11th December 1997. The Kyoto Protocol sets emissions targets for developed countries, during the first commitment period (2008-2012), which are binding under international law (United Nations, 1998).

The Kyoto Protocol features a variety of mechanisms which a participating country can use to meet its reduction targets, notably the first international emissions trading scheme. Under the Kyoto Protocol a country can purchase Kyoto Units on the international market. Carbon credits may come in the form of (i) Kyoto emissions units allocated to a country that they did not use[1] (ii) certified emission reductions units (CERs), acquired from clean development mechanisms (CDMs) (iii) Emissions reductions units, acquired from Joint Implementation (JI) schemes. All units are tradeable in an international market system (UNFCCC, 2008). An international agreement which is legally binding is desirable as the abatement of emissions provides a global benefit while the costs are levied at a lower level (producer, national) making the abatement of emissions unlikely in the absence of an agreement like the Kyoto Protocol (Bohringer & Vogt, 2003). The structure of the Kyoto Protocol is criticised for its inability to provide any meaningful reductions in GHG emissions (Copeland & Taylor, 2005; Zhang & Wang, 2011). However, the Kyoto Protocol does receive praise for being a step in the right direction, especially when considering the large geopolitical obstacles in attempting to solve a global problem (Barrett, 1998; Grubb, 2000).

The causal effects of the Kyoto Protocol on CO_2 emissions are covered thoroughly in international literature, but no direct study related to New Zealand currently exists. Previous literature provides conflicting results, dependent on which empirical method is employed. It is shown that, when countries facing emissions reductions targets are pooled together to estimate aggregate effects, the Kyoto Protocol has a significant negative effect on CO_2 emissions of approximately 7% (Grunewald & Martinez-Zarzoso, 2016; Kim et al., 2020; Maamoun, 2019). However, when estimating the effects of the Kyoto Protocol on a country-by-country basis there is no statistically significant evidence for a reduction in CO_2 emissions. When evaluating the variables used in previous causal analysis literature to

[1] Some countries were designated emissions targets above their 1990 levels.

estimate CO_2 emissions the absence of variables widely used in CGE modelling to forecast GHG emissions is absent.

While it is easy to observe that gross GHG emissions have increased relative to 1990, it is entirely possible that without the Kyoto Protocol we would be experiencing even higher global GHG emissions. Therefore, the Kyoto Protocol may have had GHG emissions reducing effects despite the “spirit” of its goals being met in a superficial manner. The importance of fighting climate change and the ambiguity surrounding how effective the Kyoto Protocol has been in reducing GHG emissions is the primary motivation behind my study. I seek to evaluate the causal effect, on New Zealand’s CO_2 emissions, of legally binding emissions reduction targets under the Kyoto Protocol. Formally, I seek to answer the following question:

Did the legally binding emissions reductions targets of the Kyoto Protocol reduce the carbon dioxide emissions of New Zealand?

The contribution of my thesis is twofold. First, I add to the Kyoto Protocol causal analysis literature with an updated data set which includes variables used in CGE modelling. Second, I provide the first causal analysis of the effects of the Kyoto Protocol on the GHG emissions of New Zealand.

Similar to Almer and Winkler (2017), my results from a Synthetic Control Matching (SCM) approach shows no evidence that the Kyoto Protocol reduced the CO_2 emissions of New Zealand. Additionally, I show that the inclusion of CGE variables within the SCM model provides an improved estimate of the counterfactual – no-Kyoto Protocol – scenario in all specifications I test. The results provide little justification for promoting the Kyoto Protocol as successful in reducing the GHG emissions of New Zealand. Rather, my thesis suggests that the only benefit of implementing the Kyoto Protocol is the “step in the right direction” argument. Furthermore, my thesis suggests that the inclusion of CGE variables consistently provides improvement to models when estimating causal effects of climate policy and therefore should be used whenever applicable.

The remainder of my thesis proceeds as follows: Section 2 provides a background on climate change, the global response, and details specific to New Zealand; Section 3 provides a comprehensive literature review which details how the environmental economic literature developed over time; Section 4 details the empirical identification strategy; Section 5 describes the data; Section 6 presents the results and robustness checks; Section 7 provides a discussion of the results; and Section 8 concludes.

Section 2 Background

In this section I provide a brief history on the development of international climate policy. It is important to have a thorough understanding of the events surrounding the Kyoto Protocol because, as outline in the methodology section bellow, international climate policies provide many obstacles when performing causal analysis. The following discussion is broken into four sections. First, I give a brief overview of the scientific consensus on climate change. Second, I describe the history of international agreements leading up to the Kyoto Protocol. Third, I describe the details of the Kyoto Protocol. Finally, I provide some context for how New Zealand approached both: the Kyoto Protocol; and international climate agreements in general. The overall timeline, as relevant to my thesis, is summarized in Table 1 at the end of the section.

2.1 Climate Change

The global climate is changing as a result of the increased GHG emissions from human activity. While surveys suggest that the public believes that climate change is a topic of scientific disagreement, approximately 97% of climate scientists have concluded that human-caused climate change is occurring (Molina & McCarthy, 2014). Climate change is defined as complex shifts, driven by GHG emissions, that affect the planets weather and climate system; the result of which will have adverse effects on the welfare of humans and wildlife (Nunez, 2019). A GHG is any gaseous compound in the atmosphere that absorbs infrared radiation, trapping and holding heat in the atmosphere. The term GHG encompasses many gasses with the three largest contributors to the increased GHG concentrations from human activity in the atmosphere being carbon dioxide, methane and nitrous oxide (Means & Lallanilla, 2021; U.S. Environmental Protection Agency, 2015).[2] The effect humans are having on the environment is accelerating, global carbon emissions have increased from 2 billion tonnes in 1900, to over 36 billion tonnes in 2015 (Ritchie & Roser, 2020). The increased emissions have resulted in atmospheric carbon dioxide concentrations being at their highest levels in over 800,000 years. With half of all man-made extra carbon dioxide, currently in the atmosphere, being placed there after 1990 (The Economist, 2019).

[2] Carbon dioxide is by far the largest contributor to GHG emissions, accounting for 76% of all GHG emissions in 2010. Methane and nitrous oxide account for 16% and 6% respectively during the same period (Intergovernmental Panel on Climate Change, 2014).

2.2 UNFCCC

The issue of climate change was first discussed at the global level during the first World Climate Conference in 1979. This set off a series of scientific and political conferences. The discourse sent out specific messages, primarily regarding the seriousness of the problem of climate change, initial targets to be adopted, and ideas of how responsibilities should be shared among countries.

Formal international unification of climate policy made its first big step forward in 1992 at the United Nations Conference on Environment and Development, informally known as the Earth Summit. During the Earth Summit the United Nations Framework Convention on Climate Change (UNFCCC) was formed. The UNFCCC is an international environmental treaty addressing climate change, negotiated, and signed by 154 countries. The primary objective of the UNFCCC, outlined in article 2 of the convention, is to stabilize the concentration of greenhouse gasses in the atmosphere "at a level that would prevent dangerous anthropogenic interference with the climate system" (United Nations, 1992).

UNFCCC acknowledges the differential role of individual countries in causing the climate change problem. Subsequently, it calls for differentiated responsibilities. The implication being that developed countries lead in reducing their own GHG emissions while assisting developing countries in both: adopting technologies that will reduce their rate of emissions growth; and adapting to climate change (Gupta, 2010). This results in countries being categorized into three main groups: Annex I, Annex II and Non-Annex I.

Non-Annex I consist of developing countries or countries which are disproportionately influenced by climate policy. Some countries are recognized as being especially vulnerable to the adverse impacts of climate change, such as countries with low-lying coastal areas and those prone to desertification and drought. Others are seen as more vulnerable to the economic impacts of climate response measures, such as countries that rely heavily on income from fossil fuel production.

Annex I include the industrialized countries that were members of the Organization for Economic Co-operation and Development (OECD) in 1992 and countries with economies in transition (the EIT parties). EITs included the Russian Federation, the Baltic States, and several Central and Eastern European states.

Annex II consist of the OECD members of Annex I, but not the EIT countries. They are required to provide financial resources to enable developing countries to undertake emissions reduction activities under the UNFCCC and to help them adapt to adverse effects of climate change. This distinction between Annex I and Annex II is to primarily account for the unique circumstances present due to the dissolution of the Soviet Union in 1991. Former soviet nations had severely underperforming economies. This negatively impacted their ability to reduce emissions and their ability to provide financial support to Non-Annex I countries (Golinski, n.d.; Gupta, 2010).

2.3 Kyoto Agreement

UNFCCC countries meet annually at conferences of the parties (COPs) to discuss how to meet the aims of the treaty. At COP-1, in Berlin 1995, the aim of Annex I countries stabilizing their emissions at 1990 levels by the year 2000 was established as "not adequate". Further discussions in the proceeding conferences directly resulted in the formation of the Kyoto Protocol at COP-3 in 1997. The Kyoto Protocol is an international treaty designed to reduce emissions of GHGs following the main principals agreed in the original 1992 UNFCCC (Grubb, 2004).

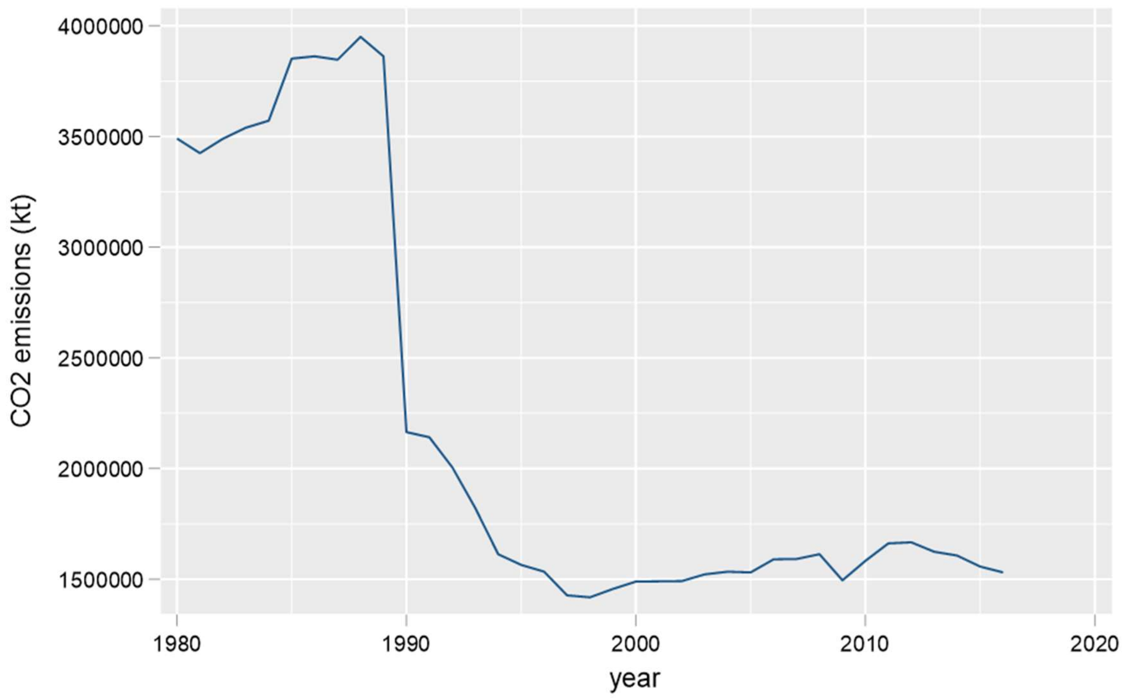
Countries are assigned a level of GHG emissions they are expected to meet, known as an emissions reduction target. These targets are given as a percentage of emissions relative to the countries own GHG emissions in 1990, for example a country given a target of 0% must limit their GHG emissions to the amount that country emitted in 1990. Six categories of GHGs are covered under the Kyoto Protocol: carbon dioxide (CO_2), methane (CH4), nitrous oxide (N20), hydrofluorocarbons (HFCs), perfluorocarbons (PFCs) and Sulphur hexafluoride (SF6). GHGs covered by the Kyoto Protocol are translated into CO_2 equivalents in determining reductions in emissions (Gielen & Kram, 1998). With the collective emissions reduction for all participating countries equating to 5%. Emissions reductions targets are specified for the period 2008-2012, known as commitment period 1 (CP1) (UNFCCC, 2008).[3]

As per article 24, the Kyoto Protocol opened for signatures from 16 March 1998 to 15 March 1999 at United Nations Headquarters in New York. The Kyoto Protocol received 84 signatures by the end of the initial signatory period (United Nations, 1998). Article 25 states that for the Kyoto Protocol to enter force no less than 55 parties to the convention, accounting for in total at least 55% of the total CO_2 emissions in 1990 must have ratified the treaty. Countries which ratify the Kyoto Protocol and face emissions reductions targets are known as Annex B countries. The ratification of Russia in December 2004 satisfied article 25 and brought the Kyoto Protocol into force, effective from the 16th of February 2005 (Gupta, 2010).

[3] To expand on the 0% reduction target example: A country given a target of 0% must limit their average annual GHG emissions to be at the same level as 1990 during the period of 2008-2012.

To meet the emissions reductions targets the Kyoto Protocol features an internationally tradeable "cap-and-trade" system of emissions trading. Countries are given assigned amount units (AAUs) that represent their GHG emissions in the year of 1990.[4] The choice of 1990, and freedom of EITs to choose their own year, as a baseline is notably advantageous to Russia. As shown in Figure 1, there is a significant drop in the CO_2 emissions of Russia from 1990 onwards. This is due their heavy industrial output, and subsequent economic decline, before the dissolution of the Soviet Union. Russia's decision to ratify the Kyoto Protocol is noted as being dependent on the choice of 1990 as a baseline year, UNFCCC countries were aware that this would results in surplus allowances under the structure of the Kyoto Protocol (Aldrich & Koerner, 2012).

Figure 1. CO_2 *emissions (kt) of Russia*



Source: Graph generated using data from the World Bank WDI database.[5]

In submitting their accounting reports for CP1 countries surrender "Kyoto Units" they are assigned (AAUs) or have acquired through the flexibility mechanisms. Certified emissions reductions (CERs) are gained through investment in developing countries (non-Annex I Parties), known as the Clean Development Mechanism (CDM). Investment in a participating

[4] Some economies in transit (EITs), primarily former Soviet Union countries, had the ability to choose a baseline date.

[5] Russia's individual annual emissions while part of the Soviet Union is calculated using their average percentage contribution to emissions of the sum of all former Soviet Union countries emissions in the 5 years after independence (1990-1994)

Annex I country would earn an emissions reductions unit (ERU), known as Joint Implementation (JI) scheme. CDMSs and JIs are meant to encourage countries to invest in economic development which reduces GHG emissions in either: Annex B countries (JI); or non-Annex B countries (CDM). This investment is rewarded with Kyoto Units, which are equivalent to the amount of emissions reductions they have generated. Reforestation/afforestation efforts and land-use change activities undertaken by a country allow the acquisition of removal units (RMU). The purpose of RMUs is to encourage the reforestation/afforestation efforts of countries, a cost effective long-run carbon sink that can aid in the stabilization of carbon content in the atmosphere (Rose & Sohngen, 2011). All Kyoto Units are equivalent in value and account for one tonne (of CO_2 equivalent) GHG emissions (UNFCCC, 2008; United Nations, 1998).[6]

While heavily criticized, the inclusion of flexibility mechanism is a diplomatic necessity required to ensure the implantation of the Kyoto Protocol (Barrett, 1998). In April 2001 the United States (US), who was to be one of the largest sources of demand for Kyoto Units, formally withdrew from the Kyoto Protocol. The exclusion of the US significantly decreases the demand in the international emissions trading scheme market, resulting in a large decrease in the cost of units, undermining the ability of the Kyoto Protocol to make significant impact to GHG emissions (Hartl, 2019). Criticism of the structure of the Kyoto Protocol is further discussed in the literature review section.

2.4 New Zealand

In terms of global GHG emissions New Zealand is a small contributor, but on a per capita basis the GHG emissions are high (*Stats NZ*, 2020). The New Zealand Government has been relatively cooperative in its dealings with international climate agreements. New Zealand signed the UNFCCC on the 4th of June 1992, the first day it opened for signatures. The subsequent Kyoto Protocol was signed on the 22nd May 1998 and ratified on the 19th December 2002 (New Zealand Ministry for the Environment, 2021).

The New Zealand Government made earnest attempts to reduce GHG emissions. In 2002 the fifth Labour government attempted to implement a carbon tax, beginning in 2007, which would apply to nearly all sectors of the economy. Notably, methane and nitrous oxide emissions from agriculture were excluded. However, Labours' coalition partners opposed the policy, and it was subsequently abandoned at the end of 2005 (Bullock, 2012). The Climate Change Response Amendment Act was introduced in September 2008, establishing

[6] Put simply GHG emissions accounting in the Kyoto Protocol is based on total GHG emissions minus any reductions in GHG emissions from investment or land use changes. Similar to a net vs gross situation.

the New Zealand Emissions Trading Scheme (NZ ETS) (Leining & Kerr, 2018). The NZ ETS was designed around the principle of compliance with the Kyoto Protocol International Emissions Trading Market. The assumption being that prices would converge toward an efficient emissions price which aligns with global GHG emissions reductions goals. This assumption did not become a reality leading to criticism of the NZETS as insufficient and at odds with the goals of international climate commitments (Leining et al., 2020; Richter & Chambers, 2014).

New Zealand's target for CP1 was to reduce average annual GHG emissions to 1990 levels. In 2015 New Zealand submitted its "True-up report" for CP1 showing that it had met its obligations (New Zealand Ministry for the Environment, 2021). However, it is important to understand how New Zealand met its obligations. As noted above, the Kyoto Protocol allows for the use of flexibility mechanisms, in conjunction with international emissions trading, to meet emissions reductions requirements. The excess allocation of emissions units to FSU countries ("hot-air") provides a low-cost method of compliance that disincentivises any real reduction in GHG emissions. New Zealand use of "hot-air" to meet its CP1 obligations is blatant when one examines the Kyoto Units submitted by New Zealand to the UNFCCC. During CP1 New Zealand produced approximately 63 million (t CO_2 equivalent) of GHG emissions above their allowed amount, but submitted approximately 96 million "hot-air" Kyoto Units purchased from Russia and Ukraine (New Zealand Ministry for the Environment, 2015).[7] While the excess retirement of purchased units poses questions, my intent is to demonstrate that it was well within the capacity of New Zealand to meet its Kyoto obligations through the purchase of international units supplied by FSU nations.

[7] More precisely New Zealand produced 63,232,888 (t CO_2 equivalent) GHG emissions in excess of their CP1 allowance and purchased a combined 96,026,848 AAU units from Ukraine and Russia. Figures are my own calculations based on data provided by the New Zealand Ministry for the Environment.

Table 1. Timeline of International Environmental Policy

Year	Event	Outcome
1979	First World Climate Conference (WCC-1)	
1988	First United Nations General Assembly Resolution on climate change held	Intergovernmental Panel on Climate Chane (IPCC) established
1992	Climate Change Convention	UNFCCC established
1995	First Conferences of the Parties (COP-1)	UNFCCC objectives deemed insufficient
1997	Third Conference of the Parties (COP-3)	Kyoto Protocol Established
2002	New Zealand ratifies the Kyoto Protocol	
2004	Russia ratifies the Kyoto Protocol	Clause 24 of the Kyoto Agreement is satisfied
2005	Kyoto Protocol Enters into Force	
2008	New Zealand Emissions trading scheme begins	
2008-2012	Kyoto Protocol First Commitment Period (CP-1)	

Section 3 Literature Review

In the following section, I provide an overview of how the Kyoto Protocol, and environmental agreements, have been explored within the literature. I provide a timeline of the evolution of the approach in the literature, which roughly coincides with the three subsections. First, I show how the literature is initially focused on theoretical explorations of the possible outcomes, based on the structure of the Kyoto Protocol agreement. There is substantial overlap between the theoretical literature and CGE modelling literature, therefore I also show how the theoretical predictions are tested withing CGE models. Second, I overview the methodology and findings of causal analysis literature on the Kyoto Protocol. Third, I outline how there is a distinct disparity between CGE and causal analysis literature in variable selection. The section concludes with my plan to fill the gap I have observed in the literature and a formalisation of my research question.

3.1 Theoretical Models

Theoretical framework, tested in a CGE model, is a common theme of the early literature on the Kyoto Protocol. Initially the focus is on the structure of the Kyoto Protocol, analysing how efficient it would be in modifying a national planners optimal abatement level. While in theory environmental economics suggests that the Kyoto Protocol would be successful, several papers predict that the theory will not hold in the real world.

Bohgringer & Vogt (2003) provide a basic framework describing a country's choice of GHG abatement levels. In the absence of an international authority, countries behave non-cooperatively, with each country deciding its GHG abatement according to a comparison of its own benefits from abatement and its own costs of abatement. Total global abatement (Q) is simply the sum of all individual countries abatement levels (q_i). GHG emissions abatement is a global public good, so individual countries benefit from abatement is a function of total global abatement. However, a countries' costs of abatement are a function of the national level of abatement that they choose. These three principles are summarised in the following:

$$Q = \sum_i q_i \quad B_i = B_i(Q) \quad C_i = C_i(q_i)$$

In a non-cooperative Nash Equilibrium, countries choose abatement levels which equate their individual marginal befits from abatement and their individual marginal abatement costs.

$$B_i' = C_i'$$

This non-cooperative Nash Equilibrium is sub optimal from the perspective of a global planner. A national planner does not account for the positive externalities experienced by all other countries as a result of their own emissions abatement. Although this simple model implies some contribution towards the public good, the authors find under the assumptions they make that optimal abatement level, and subsequent compliance cost, are close to zero.

The Kyoto Protocol is an attempt to correct this fundamental incentive issue of non-cooperation. Through the development of an international treaty with legally binding emissions targets the national planners cost of abatement is shifted. Previously, the minimum abatement cost was zero. Now, a national planner for a country facing emissions reductions targets under the Kyoto Protocol incurs a minimum cost associated with complying with Kyoto obligations.

Bohringer & Vogt (2003) criticize the structure of the Kyoto Protocol, theorising that its design is not sufficient to change the decision making of a national planner. The theory is tested in a global CGE model calibration using the Global Trade Analysis Project (GTAP) database and confirms the hypothesis that the Kyoto Protocol results in business-as-usual emissions. The paper cites the absence of the US, unrestricted emissions trading and low compliance costs as the factors which will reduce the Kyoto Protocols effectiveness at reducing global emissions.

Copeland & Taylor (2005) extend environmental economic theory by evaluating how the predictions of environmental economics fair in a world with international trade. They show that when emissions permit trade does occur, it may make both participants worse off and increases global emissions. The theoretical results are highly dependent on how permit trade, and production substitution, effects the terms of trade. There theoretical outcomes boil down to, will the production substitution effect of high emitting goods in an unconstrained country be outweighed by the income effects increasing demand for environmental quality. The Kyoto Protocol will increase the relative price of dirty goods, increasing the production of these goods in unconstrained countries, also known as carbon leakage. However, the increased income and terms of trade position will increase the countries demand for environmental quality. So, the theoretical outcome of the Kyoto Protocol depends on two factors: The prevalence of carbon leakage and how powerful the effect of income is on demand for environmental quality.

3.1.1 The income-environment relationship

The effects of income on demand for environmental quality is an area which receives considerable attention in the literature. First hypothesized by Beckerman (1972), the relationship between per capita income and environmental quality is theorised to follow an inverted U shape. In this framework environmental protection is a luxury good, GHG emissions will increase as per capita income increases, owing to the strong positive corelation between GDP and GHG emissions. When per capita income reaches a turning point, GHG emissions begin to decrease. The phenomenon is known as the Environmental Kuznets Curve (EKC). Since the 1990's sufficient empirical data on various pollutants became available through the Global Environmental Monitoring System on air and water quality (GEMS), the Toxic Release Inventory (TRI), the CO_2 emissions estimates from the Oak Ridge National Laboratory (ORNL) and the environmental data compiled by the OECD and Eurostat. These data have allowed researchers to investigate the relationship between environmental protections and income. Shahbaz and Sinha (2019) surveyed the empirical literature, published between 1991 and 2017, on environmental Kuznets curve estimation of CO_2 emissions finding that the results of EKC estimation for CO_2 emissions are inconclusive. A given paper finding a statistically significant EKC is highly dependent on the choice of contexts, time period, explanatory variables and methodological adaptation.

3.1.2 Carbon Leakage

Carbon leakage and climate policy also receives substantial attention in the literature. Surveying the empirical literature on the sensitivity of investment to environmental regulations, both internationally and within the US, it is shown that there is no significant link between investment decisions of firms and stringency of environmental regulations (Levinson, 1995).

However, Babiker (2005) predicts a different result from the Kyoto Protocol. In estimating the possible effects of the Kyoto Protocol on the relocation of energy-intensive production it is shown that significant carbon leakage will occur. The paper adopts an oligopolistic market structure, combined with increasing returns to scale production technologies, to represent the interaction between firms producing energy-intensive goods. This representation is then embedded within a multi-regional CGE model to quantify the relocation effects of the Kyoto Protocol. The results are found to depend on the representation of the market structure and how substitutable the traded energy-intensive goods are. When energy intensive goods are modelled as Heckscher-Ohlin goods, the global carbon leakage rate ranges between 50% and 130%, in which case GHG control policies in OECD countries can lead to higher global GHG emissions.

The difference in predicted carbon leakage between Babiker (2005) and the existing empirical literature can be explained from the increased trade openness. The survey on existing empirical literature accounts for research completed before 1994, with the majority occurring during the 1980's. Trade as a percentage of global GDP was approximately between 35% and 40% from 1970 through to 1994. By the time Babiker (2005) was predicting the effects of the Kyoto Protocol, global trade openness had increased by over 25% (WDI, 2021). A paper exploring the impact of trade openness on global carbon emissions provides support for this idea. In examining the top ten emitters of CO_2 in emerging countries between 1971 and 2011 it is shown that increasing trade openness results in increased CO_2 emissions for Turkey, India, China and Indonesia (Ertugrul et al., 2016). The empirical findings match Babiker's (2005) prediction, as his model suggests that most of the carbon leakage will occur in China, India, and the dynamic Asian economies.[8]

More recently Nielsen et al (2021) explore the carbon leakage present in climate agreements. This paper overviews the structure of the Paris Agreement, noting that it has insufficient provisions to account for carbon leakage. They subsequently preform a quantitative analysis on carbon leakage during the Kyoto Protocol. By examining data from 2000 to 2014 they find that carbon leakage was a feature of the Kyoto Protocol. However, carbon leakage was dominated by OECD countries which did not ratify the Kyoto Protocol, predominately occurring between the US and China. This result supports Babiker (2005) predictions, as his model examined carbon leakage from OECD countries, not from Kyoto Protocol participating countries.

3.1.3 Overview of CGE/Theory

The literature reviewed so far focuses on research produced before CP1 of the Kyoto Protocol. Theoretical work, tested in CGE models, provided a reasonably consistent narrative. The Kyoto Protocol will at best, provide no change to global emissions and at worst increase global emissions. This result is dependent on how the Kyoto Protocol works. Assuming the Kyoto Protocol is successful in increasing the abatement level chosen by a national planner. The predictions are that the Kyoto Protocol will drive high-emitting production from a Kyoto Protocol constrained developed country to a non-Kyoto Protocol constrained developing country. Firms in high-emitting production will face lower regulations and effective emissions per unit of output would increase, subsequently increasing global emissions (Babiker, 2005). While it was shown that the EKC effect may offset the carbon leakage issue in the long run, empirical evidence of an EKC is inconclusive

[8] Dynamic Asian Economies is a country grouping comprising Chinese Taipei; Hong Kong China; Indonesia; Malaysia; the Philippines; Singapore and Thailand

while the carbon leakage effect seems to find more consistent support (Nielsen et al., 2021; Shahbaz & Sinha, 2019; Stern, 2004). The design of the Kyoto Protocol was criticized with many authors noting that the flexibility mechanisms reduces the cost of emissions abatement, and would result in a business-as-usual scenario. The flexibility mechanisms mean that a national planner will not need to provide any real emissions abatement to meet their target. Even in the absence of flexibility mechanisms, the enforcement mechanisms are not strong enough to alter the national planner's decision making (Bohringer & Vogt, 2003; Copeland & Taylor, 2005).

3.2 Causal Analysis Literature

Since the ending of CP1, researchers have gained access to data which can be used to quantify the causal effects of the Kyoto Protocol on emissions. The general empirical approach to causal estimation of an environmental agreement is to compare the current GHG emissions level with the counterfactual emission levels if a given policy had not been introduced. The aim is to establish a suitable counterfactual outcome, subsequently find the business-as-usual emissions level and compare the counterfactual estimate to the observed actual outcome.

The Kyoto Protocol is an agreement which a country enters voluntarily. The voluntary nature of the Kyoto Protocol produces a selection bias issue, as the motivations of a government to join the agreement likely depend upon endogenous country specific characteristics. Selection bias must be considered when analysing the effectiveness of policies to avoid biased results (Bennear & Coglianese, 2005). Empirical investigation into the determinants of Kyoto Protocol ratification decision show that population growth, education levels, energy consumption and emissions growth are the main factors affecting the decision to ratify the Kyoto Protocol (York, 2005; Zahran et al., 2007). These determinants are mostly adhered to throughout the causal Kyoto Protocol literature, with the main difference between the research being the estimation method employed.

In cases like the Kyoto Protocol, where different groups are either exposed or not exposed to a certain treatment over a given time period, the difference-in-differences (DiD) approach is the most common method applied to overcome the issue of selection bias. The issue with DiD is that it assumes unobservable differences are constant over time, which in the case of environmental policy is an imprecise assumption, owing to the presences of cofounders which vary over time.

Grunewald & Martinez-Zarzoso (2016) use matching before the DiD estimation, to increase the similarity of the control and treated group, thus limiting the effects of the unobservable time varying cofounders. Their paper tests the effect of ratifying the Kyoto Protocol on aggregate CO_2 emissions of the industrialised countries that are legally bound by the protocol. Propensity scores of the ratification of the Kyoto Protocol are used to match treated countries to the closest country that is not treated. Propensity scores seek to reduce the unobservable differences between the treated and control group, thus reducing the bias that would result from the lack of parallel trends between the countries. The authors use data ranging from 1992-2009 on the growth rates, and squared growth rates, of CO_2 emissions, population and GDP to calculate the propensity scores of ratifying the Kyoto Protocol. Nearest neighbour matching is used to create the control group which is then used for the DiD estimation. Their results show a statistically significant 7% reduction in CO_2 emissions for countries that are facing legally binding Kyoto Protocol commitments. The mean values of the predictors used in the matched control group are shown to be sufficiently close to that of the treated group while also showing parallel emissions trends in the pre-treatment period.

Kim et al (Kim et al., 2020) reproduced the Grunewald & Martinez-Zarzoso (2016) study with an amended data set. Measures of human capital, labour force participation, capital formation and oil pricing are included. The authors also extend the study by including two possible treatment dates: 1997 and 2005. The former is chosen to represent the international agreement as treatment, while the latter represents the agreement becoming legally binding. Their results show that, using 2005 as treatment, CO_2 emissions are reduced by 13% more in countries facing legally binding emissions. However, when using 1997 as a treatment date the results are statistically insignificant.

It in their causal analysis of the Kyoto Protocol, Almer and Winkler (2017) consider the Synthetic Control Method (SCM) developed by Abadie et al (2010) to be a more feasible method. The SCM formulates a synthetic counterfactual that behaves near identically to the original treated country in the years prior to treatment, providing an accurate comparison between a treated country and the hypothetical outcome in an untreated scenario. The author's chosen variables are similar to the other causal analysis literature I discuss; including economic and societal factors which are predictors of CO_2 emissions and selection into treatment. As is required with the SCM, the data spans a far larger range than any other causal literature reviewed, ranging from 1980 to 2012. This is to ensure that the synthetic control matches the treated unit for a long pre-treatment period, minimising the likelihood of bias arising from unobservable cofounders.[9] Where the other methods reviewed provide causal estimates of the Kyoto Protocol for an aggregate treated group, this study analyses each treated country individually providing causal effects on a country-by-country basis. Treatment effects are calculated using the year a given country ratifies the Kyoto Protocol

[9] Further discussed in Methodology section bellow.

as the treatment date. They found no statistically significant evidence that supports the Kyoto Protocol effecting CO_2 emissions. Furthermore, they show that when forming a control group US state level data is preferable to country level data. The authors attribute this to differences in unobservable cofounders being comparatively smaller in the US state control group.

Seeking to combine the previously discussed methods, Maamoun (2019) employs the Generalised Synthetic Control Method (GSCM). The GSCM allows the aggregation of the treated units while also accounting for different treatment periods for each unit composing the aggregate treated group. A limitation of the method, as noted by the authors, is that it requires a larger donor pool than the SCM. Consequently, the authors have a significantly larger donor pool of countries which include developing nations. In contrast, Almer and Winkler (2017) exclude developing nations from the donor pool to increase the comparability of the treated and control units. Nonetheless, the GSCM results show a statically significant reduction in CO_2 and GHG emissions of 7% using similar variables as Almer and Winkler (2017). The results hold through a series of robustness tests while also showing that the use of US State level data is preferable when quantify treatment effects of the Kyoto Protocol.

The results of causal analysis are dependent on the methodological approach. When countries are aggregated into Kyoto and no-Kyoto observations, it is shown that the Kyoto Protocol has a statistically significant effect in reducing CO_2 emissions of approximately 7%. However, when attempting to estimate the effects on a country-by-country basis there is little statistically significant evidence for the Kyoto Protocol reducing CO_2 emissions. The causal literature is summarised in Table 2.

Table 2. Summary of the Causal Analysis Literature

Author	Method	Results
Grunewald & Martinez-Zarzoso (2016)	Difference in Differences Propensity Score Matching	7% reduction in CO_2 emissions
Almer & Winkler (2017)	Synthetic Control Method	Small, but not statistically significant, reductions in emissions
Maamoun (2019)	Generalised Synthetic Control Method	7% reduction in CO_2 emissions
Kim et al (2020)	Difference in Differences Propensity Score Matching	Up to a 14% reduction in CO_2 emissions

3.3 Variable Choice in Causal Analysis and CGE literature

While there exists some similarities, there are some notable differences in the variable selection between causal and CGE literature. Although my thesis is a causal analysis, I believe there has been too little attention placed on the insights provided by CGE models. As outlined by Sue Wing (2009), CGE models can be used to assess the impacts of climate policies on the reduction of GHG emissions. It is therefore reasonable to assume that CGE models have a relatively robust set of predictor variables employed to model emissions. For an overview of more recent applications of CGE modelling to climate policy readers should consult studies in the 36th Energy Modelling Forum study summarised by Böhringer et al (2021). The differences between causal analysis and CGE literature in variable selection for predicting emissions is summarised in Table 3

Table 3. Variables used to predict/model emissions

Causal Literature	CGE Literature
GDP	GDP
Emissions trends	Emissions trends
Population Growth	Population Growth
Sector Contribution to GDP	Energy generation mix
Human Capital Index	Sector Contribution to GDP
Life Expectancy	Agricultural land use
Political Rights	Technological progress
Oil Pricing	

Notes: CGE variables are based on COFFEE-TEA, an integrated assessment model framework

Note the difference between the causal analysis literature and the CGE modelling literature when choosing predictors of emissions. Both overlap in some areas; both include GDP, emissions trends and population data. However, the causal literature focuses on variables which control for selection bias and ignores variable which are necessary in CGE modelling. Most obvious is the lack of energy sector variables in the causal literature. In a review of 17 state of the art recursive dynamic CGE models it is shown that incorporating energy sector characteristics is a best practice when modelling baselines and alternative scenarios of GHG emissions (Faehn et al., 2020). Land use is another characteristic outlined as crucial in CGE modelling which is noticeably absent from causal analysis. The justification for including agricultural land use is that the majority of GHG emissions from agricultural activities are not directly energy-related, consisting of methane and nitrous oxide along with changes in carbon sequestration in agricultural land and forestry. Accounting for agricultural characteristic is especially relevant in my case given the large, relative to other developed OECD nations, agricultural sector of New Zealand (Andrew & Forgie, 2008).

There appears to be a disconnect in the variable choice of CGE modelling and causal analysis literature, with no fundamental reason for common CGE variables to be excluded within a causal analysis of the Kyoto Protocol. It is my intent to produce the first causal analysis of the Kyoto Protocol on the CO_2 emissions of New Zealand, considering the best practices of both methods. In doing so I plan to fill in a gap in the literature for both: the causal effects of the Kyoto Protocol on New Zealand's CO_2 emissions; and the impact of energy sector/land use characteristics in causal analysis of climate policy. More specifically I aim to answer the following research questions:

Did the legally binding emissions reductions targets of the Kyoto Protocol reduce the carbon dioxide emissions of New Zealand?

Are causal analysis specifications for the Kyoto protocol improved be including variables used to forecast emissions in computable general equilibrium modelling?

The choice of CO_2 emissions in place of GHG emissions is discussed further in the data section bellow.

Section 4 Methodology

The objective of this thesis is to identify the causal effect of the Kyoto Protocol on New Zealand's CO_2 emissions. The estimation strategy is important as I need to establish how New Zealand's CO_2 emissions would have developed in the absence of the Kyoto Protocol. In the following, I explain my strategy for estimating the causal effects in four parts. First, I outline Rubin's Model for Causal Inference. Second, I describe in detail my method for estimating the counterfactual outcome, the SCM. Third, I explain the methods for statistical inference. I finalize this section with a description of the necessary contextual requirements when using the SCM.

4.1 Causal Estimation

The main challenge for estimating a causal effect is a missing data problem, I cannot observe a country being exposed to the Kyoto Protocol and not exposed to the Kyoto Protocol at the same time. In my context, I observe the CO_2 emissions for New Zealand when they are subjected to an emissions reductions target under the Kyoto Protocol but cannot observe the counterfactual outcome in which New Zealand does not face an emissions reductions target. The causal effect of the Kyoto Protocol is simply the difference between the observed emissions and the counterfactual emissions.[10] Given that actual CO_2 emissions are observed, estimating the causal effect of the Kyoto Protocol on CO_2 emissions is equivalent to estimating the counterfactual outcome.

My research question is expressed using notation from Rubin's Model for causal inference (Rubin, 1974). Consider a given country (j), that either faces or does not face Kyoto obligations (D_j) and produces a given level of CO_2 emissions (Y_j). The treatment status of the country is described by the following:

$$D_j = \begin{cases} 1 & Country\ faces\ Kyoto\ obligations \\ 0 & Country\ does\ not\ face\ Kyoto\ obligations \end{cases}$$

The objective of the Kyoto Protocol is to reduce the CO_2 emissions of participating countries. Letting Y_j represent the GHG emissions of country j, it follows that:

$$Y_j(D_j) = \begin{cases} 1 & \mathrm{CO_2}\ emissions\ when\ facing\ Kyoto\ obligations \\ 0 & \mathrm{CO_2}\ emissions\ when\ not\ facing\ Kyoto\ obligations \end{cases}$$

[10] I refer to "legally binding emissions reductions targets under the Kyoto Protocol" as "Kyoto obligations" interchangeably.

Therefore, the observed outcome for a given country is written as:

$$Y_j = D_j Y_j(1) + \left(1 - D_j\right) Y_j(0) \tag{1}$$

Note that equation (1) implies that if a country faces Kyoto obligations ($D_i = 1$) then the outcome when a country does not face Kyoto obligations $\left(Y_j(0)\right)$ disappears and we are left with the outcome when a country is facing Kyoto obligations, vice versa. To keep notation concise, let Y_j^I denote the CO_2 emissions of a country j which faces Kyoto obligations and Y_j^N denote the CO_2 emissions of a country j which does not face Kyoto obligations.[11] Climate change policy has effects which are not instant, therefore it is important to consider how Kyoto obligations effects the CO_2 emissions of a country over time. For a country ($j = 1$) which is affected by Kyoto obligations at time period T_0, and a post-treatment period ($\mathrm{t} > \mathrm{T}_0$), the response of CO_2 emissions to Kyoto obligations is defined as $\mathbf{Y}_{\mathbf{1}t}^{\mathbf{I}}$ and the response of CO_2 emissions without Kyoto obligations is defined by $\mathbf{Y}_{\mathbf{1}t}^{\mathbf{N}}$. The causal effect of Kyoto obligations for country one in time period t (with $\mathrm{t} > \mathrm{T}_0$) is formalised in equation (2):

$$\alpha_{1\mathrm{t}} = \mathrm{Y}_{1t}^{\mathrm{I}} - \mathrm{Y}_{1t}^{\mathrm{N}} \tag{2}$$

Because country one faces Kyoto obligations at T_0 it follows from equation (2) that for $\mathrm{t} > \mathrm{T}_0$ we have $\mathbf{Y}_{\mathbf{1}t} = \mathbf{Y}_{\mathbf{1}t}^{\mathbf{I}}$. This is a statement, in terms of the framework, that for a country subjected to Kyoto obligations and a post-treatment period, the potential outcome under Kyoto obligations is observed. The challenge for my thesis is estimating how the CO_2 emissions for country one would have developed in the absence of Kyoto obligations, $\mathrm{Y}_{1t}^{\mathrm{N}}$ for $\mathrm{t} > \mathrm{T}_0$, the counterfactual outcome.

To estimate the counterfactual outcome, it is necessary that I can observe other aggregate entities that are comparable to New Zealand on relevant factors, and whose CO_2 emissions are not affected by the Kyoto Protocol. I use a comparison group that contains non-Annex B countries and the United States (US). I consider the US as not exposed to the Kyoto Protocol as it did not ratify the Kyoto Protocol so never faced binding emissions targets[12]. In following previous literature on the causal effects of the Kyoto Protocol, I also estimate a second

[11] I is used a reference to intervention. Likewise, N refers to no intervention. The word "intervention" is used interchangeably with "treatment" throughout my thesis.

[12] For the remained of the thesis I will refer to the control group as "non-Annex B countries" as this more concise than "non-Annex B countries plus the US"

specification that consists of individual state level data for the US (Almer & Winkler, 2017; Maamoun, 2019).

There are several potential strategies for estimating the counterfactual outcome when different groups are either exposed or not exposed to an intervention over a certain period (Imbens & Wooldridge, 2009). In cases like mine, which involve a single treated unit, the Difference-in-differences (DiD) approach is the most common strategy employed (Bertrand et al., 2004). A DiD analysis would involve finding a large aggregate entity which shares similar characteristics to New Zealand and was not exposed to the Kyoto Protocol. While the DiD model is valid in the right context, in practice it is often difficult to find a single (or group of) country which is suitable for use as a control group.

To overcome this limitation, I propose the use of the Synthetic Control Method (SCM), first introduced by Abadie and Gardeazabal (2003). The underlying idea of the SCM is that a linear combination of countries not exposed to the Kyoto Protocol is a better comparison group for New Zealand than any single country. Rather than a researcher choosing a single (or group of) country as a control, the SCM identifies an "optimal" comparison group by minimizing the pre-treatment difference with New Zealand for a given set of relevant emissions characteristics. This creates a "Synthetic New Zealand" which serves as the comparison group.

4.2 The Synthetic Control Method

The proceeding description of the SCM closely follows the recently published paper by Abadie (2021) which "…aims to provide practical guidance to researchers employing synthetic control methods." However, to keep my section thorough I reference the original sources which comprise the recent paper.

Consider there are $J+1$ countries: $j = 1,2,\dots,J+1$, where $j=1$ denotes the treated country, in my case this is New Zealand facing binding emissions target under Kyoto Protocol, and $j = 2,\dots,J+1$ is all untreated countries, the control group. The data spans T periods with: T_0 being the period of treatment; $t < T_0$ being the periods before the treatment; and $t > T_0$ being the periods after the treatment. It follows that for each unit (j), and time (t), we observe the CO_2 emissions (Y_{jt}). For each unit (j), we also observe a set of predictors (k) of GHG emissions ($X_{1j},\dots,X_{Kj}$) which are unaffected by the Kyoto Protocol.[13] The $k \times 1$ vectors $\boldsymbol{X}_1,\dots,\boldsymbol{X}_{J+1}$ contain the values of the predictors for countries

[13] Predictors may include pre intervention values of CO_2 emissions as these values have not been affected by treatment. It is relatively common in literature using the SCM to include some pre intervention values of the outcome of interest. However, it is also bad practice to include all pre intervention values of the outcome as this leads to substantial bias in the estimator (Kaul et al., 2015).

$j = 1, \ldots, J + 1$. The $k \times J$ matrix $\boldsymbol{X}_0 = [\boldsymbol{X}_2 \ldots \boldsymbol{X}_{J+1}]$, contains the values of the predictors for all countries the donor pool. As noted above, I have the data on the actual CO_2 emissions path ($\mathrm{Y}_{1t}^{\mathrm{I}}$) for New Zealand while I do not have information about the counter factual emissions which would have occurred if New Zealand had not been exposed to the intervention (i.e., $\mathrm{Y}_{1t}^{\mathrm{N}}$ for $t > T_0$). Abadie (2021) proposes the use of observed characteristics of the countries in the donor pool to form a synthetic control.[14] Using this method, Synthetic New Zealand is represented by a $J \times 1$ vector of weights $W = (\omega_2, \ldots, \omega_{\mathrm{J}+1})'$. Given a set of weights, W, the counterfactual outcome $\mathrm{Y}_{1t}^{\mathrm{N}}$ is given by:

$$\mathbf{Y}_{\mathbf{1}t}^{\mathbf{N}} = \sum_{j=2}^{J+1} \omega_j Y_{jt} \tag{3}$$

and therefore, the treatment effect $\boldsymbol{\alpha}_{\mathbf{1t}}$ is:

$$\boldsymbol{\alpha}_{\mathbf{1t}} = \mathbf{Y}_{\mathbf{1}t}^{\mathbf{I}} - \sum_{j=2}^{J+1} \omega_j Y_{jt} \tag{4}$$

4.2.1 Determining Weights

Choosing W weights is a simple process at a surface level. Suitable comparison units are placed into a donor pool, a formalised data driven process is used to give each donor pool country a weight (ranging from 0 to1) with the sum of all weights equalling one. The weights W will be chosen so that the weighted average of all countries in the donor pool resembles New Zealand with respects to CO_2 emissions in the pre-intervention period and all other relevant characteristics ($\boldsymbol{X}_1$).[15] Formally, I seek W such that for all $t < T_0$:

$$\sum_{j=2}^{J+1} \omega_j^* Y_{jt} = Y_{1t} \tag{5}$$

And

$$\sum_{j=2}^{J+1} \omega_j^* \boldsymbol{X}_0 = \boldsymbol{X}_1 \tag{6}$$

[14] A donor pool comprises a set of suitable comparison units.

[15] Note that countries can be interchanged with units, states, entities etc. My use of the word countries is motivated by the country level data being the primary specification.

In constructing a suitable synthetic control, I need to determine a set of weights $W = (\omega_2, \dots, \omega_{J+1})'$, subject to two constraints. First, any given weight must be non-negative, $\omega_j \geq 0$. Second, all assigned weights must sum to one, $\sum_{j=2}^{J+1} \omega_j = 1$. Given the non-negative no extrapolation conditions, a vector W such that equations (5) and (6) hold may not exist. In an ideal world, there would exist such a combination of countries which would satisfy equations (5) and (6) but the restrictive nature of requiring absolute parity makes this unrealistic. Therefore, as a compromise I will choose weights subject to the following distance equation:

$$||X_1 - X_0 W|| = \left(\sum_{h=1}^{k} v_h \, (X_{h1} - w_2 X_{h2} - \dots - w_{J+1} X_{hJ+1})^2 \right)^{1/2} \tag{7}$$

Instead of requiring complete parity between an estimated synthetic control and the treated country, I simply look to minimise the differences in my chosen predictor variables. However, not all predictor variables have equal power in predicting emissions. The set of constants, $v_1, \dots, v_k$, are non-negative and reflect the relative importance of the synthetic control reproducing the values of the k predictors for New Zealand, $X_{11}, \dots, X_{k1}$. Consequently, the optimal weights, W, depend on the weighting of matrix $V = (v_1, \dots, v_k)$. For each potential V there is a corresponding synthetic control $W(V) = \left(w_2(V), \dots, w_{J+1}(V)\right)'$, which is calculated by minimizing equation (7). The next logical step is determining V , however there is little guidance in the literature as to the best method of choosing V weights. In their study on the Kyoto Protocol, Almer and Winkler (2017) opted to calculate W weights using even V weights. Abadie, Diamond and Hainmueller (2010) suggest choosing V such that the synthetic control minimizes the mean squared prediction error (MSPE) of the synthetic control, for a given set of pre-treatment periods, with respects to Y_{1t}^N:

$$\sum_{t \in T_0} (Y_{1t} - w_2(V) Y_{2t} - \dots - w_{J+1}(V) Y_{J+1t})^2 \tag{8}$$

Note that all estimation equations so far have occurred during the pre-treatment period. By this stage in an estimation, I will have established the relative importance that each predictor variable has for replicating the CO_2 emissions of New Zealand and subsequently found my W weights. However, if the V weights are calculated using the entire pre-treatment period it is uncertain if the subsequent synthetic control is a suitable predictor of New Zealand's CO_2 emissions for anything other than the pre-treatment period. In order to overcome this, I use out-of-sample validation in the estimation stage of the equation. I divide the pre-treatment periods into a training and validation period. The first half of the pre-treatment periods will be the training period in which I minimise the MSPE of equation (8) to obtain the optimal V weights, V^*. The final half of the pre-treatment period is the validation period where I will use $V^*$and data on the predictors of emissions to calculate the synthetic control $W^* = W(V^*)$. After I have calculated a suitable synthetic control, I can estimate the treatment effect in equation (2) for any given time post-treatment time period.

4.3 Significance Tests

By design the SCM does not produce a significance test on the treatment effect. This is a problem as I may have produced a 'Synthetic New Zealand' which closely tracks the CO_2 emissions of New Zealand in the pre-treatment period but have no way of evaluating the probability that estimated post-treatment outcomes happened by chance.

However, Abadie, Diamond and Hainmueller (2010) propose the use of permutation methods or placebos tests. In following the placebo test approach, I iteratively assign treatment to units in the donor pool, estimating "placebo effects" of treatment for each country. I then plot the estimated placebo effects along with the estimated treatment effect for New Zealand. As the donor pool countries have not been exposed to treatment, I should see no effect from treatment. The treatment effect of the Kyoto Protocol on the GHG emissions of New Zealand can be deemed significant when its magnitude is extreme relative to the placebo distributions.

The placebo test method is not always reliable. For example, imagine, given a certain specification, we find that Synthetic New Zealand closely matches the pre-treatment trajectory of New Zealand's CO_2 emissions but fails to do so for the units of the donor pool. Following this the treatment effects estimated for the placebos are also larger for the placebos. In cases like this it is preferably to use the ratio of the post-treatment fit relative to pre-treatment, known as the root mean squared prediction error (RMSPE), as a test statistic (Abadie et al., 2010).

$$R_j(t_1, t_2) = \left(\frac{1}{t_2 - t_1 + 1} \sum_{t=t_1}^{t_2} (Y_{jt} - \hat{Y}_{jt}^N)^2 \right)^{\frac{1}{2}} \tag{9}$$

The RMSPE estimations can then be used to calculate a pseudo p-value using the following equation:

$$p = \frac{1}{J+1} \sum_{j=1}^{J+1} I + (r_j - r_1) \tag{10}$$

As a small example imagine I estimate the effects of the Kyoto Protocol on CO_2 emissions of New Zealand using a donor pool consisting of 19 countries. Assume that the estimated treatment effect for New Zealand is larger than the subsequent placebo tests. The pseudo p-value would therefore be: $p = \frac{1}{20} = 0.05$. Placebo tests can provide a useful tool for analysing the probability of obtaining a treatment effect by chance, but consideration must be given to the context of the research, and its subsequent results in the discussion. While there is progress to be made in the development of alternative statistical inference testing method for the SCM, this is beyond the scope of my thesis. I choose to rely on the tools outlined by Abadie (2021) and give the relevant considerations in my discussion section.

4.4 Contextual Requirements

For the Synthetic Control Method to be appropriate there are a set of contextual requirements that must be met, and robustness checks which must be passed, to minimise the likelihood of biased results. In the following I give an overview of these requirements (robustness checks), and why they are important to establish an unbiased synthetic control.

4.4.1 Size of effect and volatility of the outcome

In estimating the effect of policy interventions with the SCM there must be a sizeable effect relative to the volatility of the outcome variable. For example, if the reduction effects to CO_2 emissions are small, the effects of the Kyoto Protocol will be indistinguishable from other exogenous shocks to CO_2 emissions. Likewise, if CO_2 emissions are too volatile, then even a large treatment effect will be difficult to distinguish from the intrinsic volatility. However, volatility issues primarily arise from volatility caused by country-specific factors. If the volatility in CO_2 emissions arise from a common factor, it can be controlled for by choosing an appropriate specification.

4.4.2 Relevant Comparison Group Available

There must be a group of units available which have not been exposed to the Kyoto Protocol, or any treatments similar to it, to make up the donor pool. Units within the donor pool must also have not experienced any idiosyncratic shocks to their CO_2 emissions, within the time period used to evaluate the Kyoto Protocol, if it is reasonably assumed that these shocks would not have affected the CO_2 emissions in the absence of the Kyoto Protocol. I must ensure that countries in the donor pool share similar characteristics to New Zealand. Although the restrictions placed on the weights, W, restrict extrapolation, interpolation biases may occur if large differences in characteristics are averaged away in the matching stage. While there is no official minimum number of comparison countries, ideally, I want at least 19 so as to be able to obtain a pseudo P-value of 0.05.[16]

4.4.3 No Anticipation

As the SCM exploits time variation in the outcome variable, the estimator may be biased if forward-looking agents act in anticipation of the treatment, or if certain aspects of the treatment are prematurely installed. For example imagine a scenario in which economic agents within New Zealand, upon learning of New Zealand ratifying the Kyoto Protocol, being to reduce their CO_2 emissions behaviour. In this scenario if I choose to select the beginning of CP1 as my treatment period (T_0) my estimated Synthetic Control will incorporate this CO_2 emissions reduction behaviour into the pre-treatment outcomes, reducing the estimated effectiveness of the policy. In order to estimate the complete effect of a treatment Abadie (2021) suggests setting the treatment date to a period before and anticipation effects can be expected. A benefit of the synthetic control method is that backdating the treatment period does not bias the estimator of the treatment effect as there is no restriction on the time variation in the effect of treatment (see equation (2) and equation (3)). If I arbitrarily set a treatment period much earlier than the actual observed treatment period the estimated treatment effects will be zero, or close to zero, up until the actual treatment has an effect.

4.4.4 No Interference

No interference is the Stable Unit Treatment Value Assumption (SUTVA) of Rubin (1980). This implies that the outcome of any one of the countries CO_2 emissions should be unaffected by the treatment status of any other country. In practice, treatments may have spill over effects on countries which are not assigned to treatment. Accounting for spill over effects is a restriction which must be imposed in the design of the study. This can be achieved by omitting countries from the donor pool whose CO_2 gasses have likely been affected by the Kyoto Protocol. The omission of countries from the donor pool can potentially cause issues in forming a relevant comparison group. I want to choose countries which are reasonably similar to New Zealand in terms of their characteristics, however this can lead to selecting countries which are more likely to be subject to spill over effects from treatment. In the case that I must include countries which may have been affected by treatment the sparse weighting, inherit to the SCM, allows me to account for potential bias in the estimator in the analysis phase. Abadie (2021) suggests that so long as I am aware, and discuss, the potential direction of bias, complete elimination of any potential spill over effects is not necessary. The presence of the flexibility mechanisms (discussed in the Background section), especially the CDMs, provide possible violations of the SUTVA/no interference

---

[16] As shown in the significance test subsection. The placebo test will never give a statistically significant result (assuming 0.05 is the cut-off) if the donor pool has less than 19 units.

requirement. My strategy in accounting for possible spill over effects of CDM's is addressed in the results section.

4.4.5 Convex Hull Condition

The use of Synthetic Control estimates is based on the idea that a combination of control units can approximate the pre-treatment characteristics of the treated unit. Once I have estimated a Synthetic Control, I must check that the difference in the characteristics of New Zealand and Synthetic New Zealand are sufficiently small. It should be noted that a Synthetic Control may struggle to replicate a variable if the treated unit is a significant outlier in the value of that variable. This is an important limitation as New Zealand's large agricultural sector is unique for a developed nation. Agriculture, as percentage of GDP, for the OECD steadily decreased from 2.2% in 1997 to 1.4% in 2019 while New Zealand remained relatively stable decreasing agriculture's share of GDP from 6% to 5.6% over the same time frame (World Bank, 2020). This may not be a concern if the estimated Synthetic Control closely tracks the trajectory of New Zealand's CO_2 emissions during a hold-out validation period. However, in the presence of outlier variables Abadie (2021) advises the transformation of the variable into time differences or growth rates or normalising the variable with respects to the treatment year value.

4.5 Robustness Checks

As with any policy evaluation, it is important to establish that my results are robust to changes in specification. By design the SCM produces weightings in the donor pool which are sparse. Abadie, Diamond and Hainmueller (2015) take advantage of the sparse weighting to run a leave-out-out reanalysis to establish robustness. Given an estimated Synthetic Control, I re-run my specification leaving one donor pool unit with positive weight out at a time. If, for any given leave-one-out test, the estimated treatment effects are significantly changed, without a large change in the pre-treatment fit, I can conclude that the 'left-out' unit has an adverse impact on the treatment effect. In this case it is advisable to drop the unit entirely as I want results to be robust to the exclusion of any given country. There is no formal method, or numerical cut-off, that defines what a significant change in estimate treatment effect relative to pre-treatment fit actually is. Abadie (2021) advises researchers using the SCM to plot the results of leave-one-out tests graphically, interpreting the outputs at their own discretion, while providing the outputs for readers to confirm (or question) the conclusions being drawn.

Section 5 Data

As outlined in the methodology section, I use the SCM to create a synthetic counterfactual New Zealand through a convex combination of all units in the donor pool. The SCM requires data on CO_2 emissions, and predictors of CO_2 emissions, for New Zealand and a set of comparison units which were not affected by the Kyoto Protocol. The SCM's credibility is predicated on its ability to track the pre-treatment trajectory of the outcome variable. Too short a pre-treatment time period may produce a synthetic control which is optimized for a unique time period and therefore is a poor representation of the long run emissions trend (Abadie, 2021). I therefore choose to collect data from 1980-2016 and form two separate donor pools. The first donor pool comprises data on non-Annex B countries. The second donor pool comprises US State level data. As a secondary research question, I look to establish if the inclusion of variables commonly used in CGE modelling provide an improvement in using the SCM to estimate the causal effects of climate policy. Therefore, my chosen predictor variables of CO_2 emissions closely follow those employed by Almer and Winkler (2017). Deviations from their specification are summarised at the end of this section. The following section is broken down as follows: I begin by outlining my choice of predictor variable; next, I outline my key variables and sources of data for the non-Annex B comparison group and US state comparison group respectively; I then provide a table showing the reduction in predictor balance achieved when using the SCM; and I conclude with a table describing the specific measure of each variable.

5.1 Outcome Variable

A common method in past causal analysis is to use CO_2 emissions as a proxy for GHG emissions. The reason being GHG emissions data is only available from 1990 onwards, while data for CO_2 emissions is available back to 1980. Furthermore, there is a strong positive correlation between CO_2 emissions and GHG emissions as shown by the pairwise test below.

Table 4. Pairwise Correlation test

Variables	GHG	CO_2	Methane
GHG	1.000		
CO_2	0.983*	1.000	
	(0.00)		
Methane	0.906*	0.847*	1.000
	(0.00)	(0.00)	

**** p<0.01, ** p<0.05, * p<0.1*

Source: Own calculations based on data from the World Bank WDI

Table 4 shows the pairwise correlation results between GHG emissions, CO_2 emissions and methane emissions for all countries and time periods in my dataset. All three variables are near perfectly correlated and statistically significant to the one percent level. While I follow the method of using CO_2 as an outcome variable, I do not consider CO_2 a suitable proxy for GHG in the case of New Zealand for two reasons. First, non-CO_2 emissions account for over half of New Zealand's GHG emissions. Second, the agricultural sector, which is absent from the NZETS, accounts for approximately half of all New Zealand's GHG emissions with agricultural emissions consisting primarily of methane and nitrous oxide gasses (New Zealand & Interim Climate Change Committee, 2019).

5.2 Country level data

For the non-Annex B analysis, I use a data set compiled from a variety of open-access resources. In following trends of previous causal literature, I pick indicators of emissions that represent economic performance, sector contribution to GDP, societal and political characteristics (Almer & Winkler, 2017; Maamoun, 2019). Furthermore, I add variables which account for energy generation and agricultural land use, which are both key considerations of CGE modelling that are notably absent from previous causal analysis literature.

I primarily use data sourced from the World Bank's Open Data, specifically the World Development Indicators (WDI). I combine the data taken from the WDI with real GDP per capita and Human Capital Index data from the Penn World tables (PWT). I also include a measure of political freedom taken from the Polity IV project (*Polity IV Project: Home Page*, n.d.; *PWT 10.0*, 2021; World Bank, 2021). The WDI data is a compilation of relevant, high-quality, and internationally comparable statistics about global development. Data is compiled from a variety of sources, usually national statistics agencies, central banks, custom services and international organizations (World Bank, 2021). A key advantage of the WDI is the ability to access data for multiple countries which has been standardised to aid in cross-country comparability.[17] Due to the limited capability of the statistical systems there are issues of data availability and comparability for some of the poorest nations in the world. Furthermore, statistical methods, coverage, practices and definitions differ widely between countries. This results in a data set which may not be suitable for cross-country analysis, despite the best efforts of the World Bank. In following with the recommendations of Abadie (2021) I normalise variables, with respects to year of treatment, which are likely

[17] The World Bank notes that full comparability cannot be assured and urges that care is taken when interpreting the indicators.

to exhibit the recording bias. However, as most variables selected are either growth rates, or proportions of totals, only CO_2 emissions require normalising to control for recording inconsistency. Normalisation is calculated by dividing a variable for any given year by the value of that variable in the base year, formalised by the following:

$$Nomalised\ X = \frac{X_t}{X_{T_0}}$$

Normalising the outcome variable has the additional benefit of forcing the emissions of New Zealand and Synthetic New Zealand to be equivalent at treatment, simplifying the interpretation of the SCM model results.

5.2.1 Countries of Interest

To evaluate the improvement of included CGE modelling variables I select the same comparison group countries as Almer and Winkler (2017). This leaves me with the following list of suitable countries to include in my donor pool:

Table 5. Non-Annex B Donor Pool

Country Name		
Brazil	Botswana	Chile
China	Columbia	Dominican Republic
Ecuador	Fiji	Gabon
Jordan	South Korea	Mexico
Mauritius	Malaysia	Panama
Saudi Arabia	Singapore	Thailand
Tunisia	Turkey	United States
Venezuela	South Africa	

To reduce the likelihood of missing observations and increase the similarity of countries in the donor pool, the country-level group is restricted to countries that are classified as high income and upper middle-income countries by the World Bank. The US is considered a suitable country for the non-Annex B donor pool as it never ratified the Kyoto Protocol and therefore was never subject to legally binding emissions reductions targets. For simplicity when I refer to “non-Annex B” this will include the US unless specifically stated otherwise.

5.2.2 Country level descriptive statistics

Table 6. Country level descriptive statistics

Variable	N	Mean	Std. Dev.	Min	Max
New Zealand					
GHG emissions	11	72572.091	4177.724	66894.903	80252.282
CO_2 emissions	11	28016.364	3307.393	23910	33440
GDP growth	10	2.533	1.26	.173	4.538
Ind. value added	11	24.526	1.166	23.11	26.538
Agri. value added	11	6.576	.771	5.729	7.981
Serv. value added	11	61.772	1.466	59.57	63.905
Life expectancy	11	77.77	.978	76.434	79.146
Human Capital Index	11	3.27	.004	3.263	3.276
GDP per Capita	11	28426.959	2278.14	24986.891	32189.854
Agricultural Land	11	55.383	5.04	45.034	60.457
Political Rights	11	10	0	10	10
Population growth	11	1.194	.494	.525	1.974
Renewable Energy use	11	28.918	1.085	26.881	30.585
Non-Annex B Countries					
GHG emissions	233	784040.32	1717281.1	1574.685	6991255.2
CO_2 emissions	253	484067.95	1255181.9	750	5776410
GDP growth	230	1.839	3.842	-13.673	10.145
Ind. value added	249	33.355	9.237	17.932	56.255
Agri. value added	249	7.766	4.939	.061	23.403
Serv. value added	249	52.75	8.565	33.57	74.881
Life expectancy	252	70.063	6.056	50.232	79.039
Human Capital Index	253	2.34	.396	1.676	3.608
GDP per Capita	253	15306.134	12348.966	454.033	51529.727
Agricultural Land	252	38.487	19.565	1.164	80.888
Political Rights	253	4.316	5.954	-10	10
Population growth	253	1.67	.739	-1.474	5.476
Renewable Energy use	253	22.807	18.416	.009	83.627

Source: My own calculations based on data from WDI, PWT and Polity IV project

Table 6 is the descriptive statistics for a 11-year sub-sample period prior to and including year of treatment (1993-2003). The 11-year subsample is chosen to reflect the specification employed. When selecting predictor variables for the SCM I must also choose which pre-treatment periods of the predictors to include. To aid in assessing the contribution that including CGE modelling variables adds to causal analysis, my choice of pre-treatment periods follows that of Almer and Winkler (2017). They opt to use the averages of two 5-year periods prior to treatment for all predictor variables. In my case this is 1998-2002 and 1993-1997, taking 2003 (ratification) as the year of treatment. After estimating a Synthetic

Control, I will be presented with the mean values of the chosen predictor variables for Synthetic New Zealand. Therefore, the descriptive statistics are chosen to be comparable to the means I will obtain from my estimated Synthetic New Zealand. Additionally, the averages for two 5-year periods post treatment for life expectancy, human capital index and population growth are included. In my case this is 2004-2008 and 2009-2013.

I extend the specification by including the averages of two 5-year pre-treatment periods of Renewable Energy (percentage of total energy use) and Agricultural land use (percentage of total land). Furthermore, I collect data ranging from an extended range (1980-2016 as opposed to 1980-2011) which is a benefit in estimating the causal effects of climate policy due to the delayed effects of emissions reductions efforts.

5.3 US state level data

As an alternative specification I also consider the use of US states as a control group. Unfortunately, multiple variables, available at a country level, are not available at a state level. This limits the available predictors to GDP and population variables.

Table 7. Descriptive Statistics Mean Comparisons

Variable	US	NZ	non-Annex B
Life expectancy	76.33	77.77	69.77
Political Rights	10	10	4.058
Human Capital Index	3.553	3.27	2.291

Notes: In this case non-Annex B excludes the United States

Source: My own calculations based on data from WDI, PWT and Polity IV project

The differences in the means of all variables listed in Table 7 are larger between New Zealand and non-Annex B countries than between New Zealand and the US.[18] It is reasonable to assume that variability at a state level, not described by aggregate US means in Table 7, exists. However, both Almer and Winkler (2017) and Maamoun (2019) establish a preference for using US state level data in estimating causal effects of the Kyoto Protocol. Thus, I work within the data availability and scope of my thesis, while considering the limitations this imposes in the discussion of my results. As a result, I assume that it is less important to control for as wide a variety of variables in my US state level specification. Noteworthy is the US state specification benefiting from a larger donor pool, 51 states as opposed to 23 countries.

As noted by Almer and Winkler, variability in the classification and recording between country and state level data, especially GDP, is expected. Following the method of Almer and

[18] Note that non-Annex B excludes the US in this case.

Winkler (2017) I normalise all variables with respects to treatment year. Data for New Zealand's variables is taken directly from the non-Annex B data set described above. Data for US states stems from multiple sources. CO_2 emissions data is collected from the US Energy Information admin. GDP data is collected from the Bureau of Economic Analysis. Population data is collected from the Population Estimates Programme conducted by the US Census Bureau (U.S. Bureau of Economic Analysis, 2021; U.S. Census Bureau, 2021; U.S. Energy Information Administration, 2021).

5.3.1 US state level descriptive statistics

Table 8. US State level Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min	Max
New Zealand					
CO_2 emissions	11	28016.364	3307.393	23910	33440
Population growth	11	1.194	.494	.525	1.974
GDP growth	11	3.968	1.493	1.06	6.328
GDP per capita	11	28426.959	2278.14	24986.891	32189.854
US States					
CO_2 emissions	561	109729.98	113151.58	3920.875	717958.67
Population growth	561	1.076	.923	-1.491	6.055
GDP growth	561	5.412	2.813	-6.854	16.465
GDP per capita	561	32334.823	11925.121	17540.971	130480.28

Source: My own calculations based on data from WDI, PWT and Polity IV project

Table 8 is the descriptive statistics for a 11-year sub-sample period prior to and including year of treatment (1993-2003). The 11-year subsample is chosen to reflect the specification employed. I estimate the causal effect of the Kyoto Protocol on the CO_2 emissions of New Zealand with US state level data following the exact specification employed by Almer and Winkler. They opt to use the averages of two 5-year periods prior to treatment for all predictor variables. In my case this is 1998-2002 and 1993-1997, taking 2003 (ratification) as the year of treatment. After estimating a Synthetic Control, I will be presented with the mean values of the chosen predictor variables for Synthetic New Zealand. Therefore, the descriptive statistics are chosen to be comparable to the means I will obtain from my estimated Synthetic New Zealand. Additionally, the averages for two 5-year periods of post treatment population growth are included. In my case this is 2004-2008 and 2009-2013.

As noted above, US state level data is preferable to country level data in the literature I review. Therefore, I aim to replicate the results that Almer and Winkler (2017) would have obtained if they had chosen to estimate the causal effects of the Kyoto Protocol on the CO_2 emissions of New Zealand and compare that to my extended country level data set. With a result from US state level data (the preferred choice of previous causal analysis), I can more accurately assess the contribution that CGE variables makes in estimating causal effects of climate policy.

5.4 SCM vs Entire Donor Pool

A key benefit of using the SCM is obtaining a synthetic counterfactual which mimics actual New Zealand on both CO_2 emissions and my chosen predictor variables. As shown by column three of Table 9, the use of SCM reduces the differences in the predictor variables when compared to the means of the entire donor pool.

Table 9. Predictor Balance

Variable	Diff. original	Diff. matched	Diff. Reduction
Non-Annex B Countries			
Entire Donor Pool			
GDP per Capita	12744.536	-2,331.98	37.12
GDP growth	0.96	-0.20	27.14
Ind. value added	-8.7	-2.35	25.76
Agri. value added	-1.1	1.51	-6.15
Serv. value added	8.8	-0.74	13.09
Life expectancy	7.567	3.65	5.04
Human Capital Index	0.9	0.10	24.69
Agricultural Land	18.05	8.67	16.59
Political Rights	5.68	2.27	34.1
Population growth	-0.56	0.01	49.54
Renewable Energy use	6.19	4.49	5.86
CDM Excluded			
GDP per Capita	12272.39	-766.83	41.01
GDP growth	1.25	0.35	32.25
Ind. value added	-8.39	0.15	33.44
Agri. value added	-1.13	2.40	-18.96
Serv. value added	8.35	-1.74	10.75
Life expectancy	8.01	3.80	5.42
Human Capital Index	0.91	0.15	23.48
Agricultural Land	18.39	6.89	20.34
Political Rights	5.48	0.32	51.78
Population growth	-0.66	-0.05	55.04
US States			
GDP per Capita	-3907.86	2046.18	6.54
GDP growth	-1.44	-1.35	2.26
Population Growth	0.11	-1.59	-124.1

Notes: Table 9 shows the average differences (across countries/states and year) between New Zealand and the two donor pools (non-Annex B countries and US states) before matching (Diff. Original) and after matching (Diff. matched). Column 3 shows the percentage decrease in the predictor imbalance relative to the scale of the given variable.

Source: My own calculations based on data from WDI, PWT and Polity IV project

5.5 Variable Descriptions

The Table below summarises the descriptions and sources of the variables used in my thesis:

Table 10. Descriptive statistics details

Variable	Description	Source
Country Level		
GHG emissions	Total Greenhouse Gas emissions (kt of CO_2 equivalent)	Word Bank WDI
CO_2 emissions	Total Carbon Dioxide emissions (kt)	Word Bank WDI
GDP per Capita*	Real GDP per capita at constant national prices (in mil, 2017 US$)	PWT 10.0
GDP growth	Calculated in Stata using GDP per capita	PWT 10.0
Agri. value added	Agriculture, forestry and fishing, value added (% of GDP)	World Bank WDI
Serv. value added	Services, value added (% of GDP)	World Bank WDI
Ind. value added	Industry (including construction), value added (% of GDP)	World Bank WDI
Human Capital Index	Based on years of schooling and returns to education	PWT 10.0
Life expectancy	Life expectancy at birth, total (years)	World Bank WDI
Agricultural Land	Agricultural land use (% of land area)	Word Bank WDI
Political Rights	Scale ranging from -10 (Autocracy) to 10 (Full democracy)	Polity IV Project
Population growth	Calculated in Stata using total population data	PWT 10.0
Renewable Energy use	Renewable energy consumption (% of total final energy consumption)	World Bank WDI
US State Level		
CO_2 emissions	Total carbon dioxide emissions (kt)	US EIA
Population growth	Calculated in Stata using total population	US Census
GDP per Capita	Total GDP per capita (in mil. 2017 US$)	US BEA
GDP growth	Calculate in Stata using GDP pe Capita	US BEA

*Variable calculated with 'Real GDP using national-accounts growth rates' (RGDPNA) and population data from the PWT 10.0

Selected variables fit into two main criteria. Variables which measure economic characteristics and variables which measure societal characteristics. Economic output shares a strong positive correlation with emissions. Therefore, it is critical to include variables which account for the economic performance and structure when modelling emissions. As the Kyoto Protocol is a treatment which a country self-selects into, it is important to include characteristics which account for the structure of the society. Note that most of the Kyoto Protocol participating countries, with binding emissions reduction targets, are developed democracies. Including societal variables, such as political rights and human capital index, can help ensure that the synthetic New Zealand closely matches actual New Zealand on likelihood to self-select and reduce the present self-selection bias.

Section 6 Results

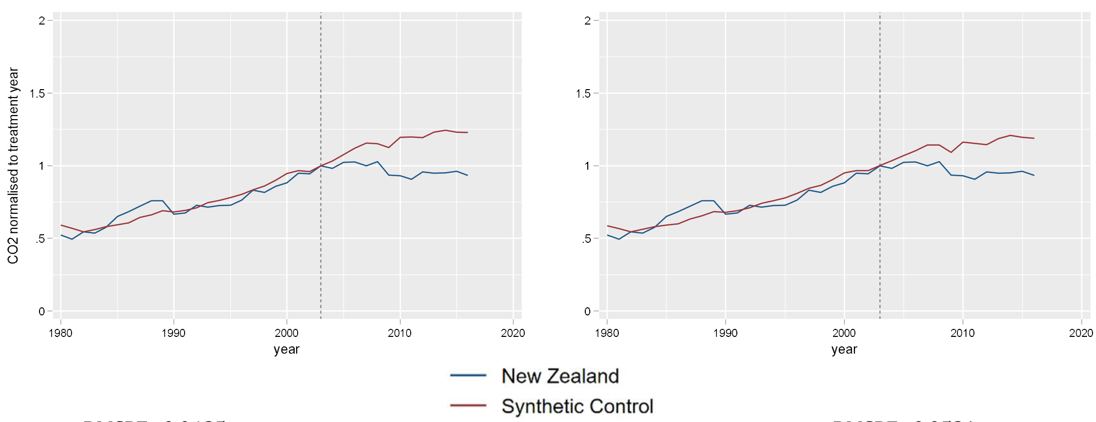
In the following section I examines the causal effects of legally binding emissions reductions targets on the CO_2 emissions of New Zealand. Section 6.1 compares the estimated effects of my specification to the effects obtained using the specification of Almer and Winkler (2017) employing a country level data set. To validate these estimates, placebo tests are produced, and subsequently the pseudo p-values are calculated. Section 6.2 restricts the country level donor pool to simultaneously control for the possibility of SUTVA violations and test the robustness of the estimates produced, for both specifications, in section 6.1. Once again, placebo tests and pseudo p-values are calculated to validate the estimates. In section 6.3 I use the leave-one out method to test the robustness of the estimates from my specification in section 6.2. In section 6.4 I run the US state level data analysis, again running placebo tests and plot the pseudo p-value.

6.1 Primary Specification

My primary specification extends a specification used by Almer and Winkler (2017) to include variables used in CGE modelling to prediction GHG emissions. The results of both my primary specification and the Almer and Winkler specification are presented visually in Figure 2.

Figure 2. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B donor pool)

a) My Specification b) Almer & Winkler Specification



RMSPE 0.0425 RMSPE 0.0521

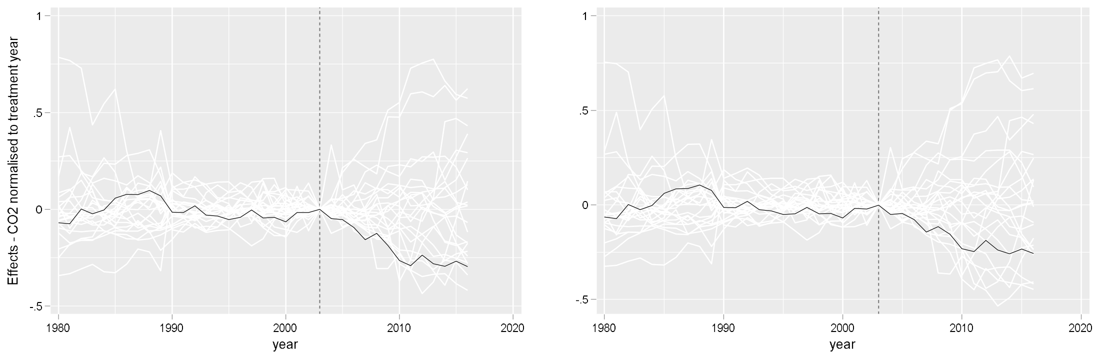
Notes: The dashed vertical line represents the treatment (ratification in 2003). The RMSPE's for the pre-treatment period are reported under each Figure respectively.

The treatment effect is given by the differences in post-treatment CO_2 emissions paths of New Zealand and Synthetic New Zealand. Both specifications provide low RMSPEs and track the pre-treatment CO_2 emissions of New Zealand closely. Although it is not obviously apparent from the figures, the lower RMSPE of my specification indicates that it produces a Synthetic Control which more accurately tracks the pre-treatment CO_2 emissions of New Zealand. Estimated treatment effects are relatively similar. My specification suggests that the average treatment effect for the entire post-treatment period is a 19.5% reduction in CO_2 emissions. The Almer and Winkler (2017) specification suggests that the average treatment effect for the entire post-treatment period is a 16.7% reduction in CO_2 emissions.

However, upon conducting placebo tests to establish the statistical significance of these results it is found that most of the estimated treatment effects, for both specifications, are not significant. The results of the placebo tests are presented visually in Figure 3 with the results of the subsequent pseudo p-value calculations being plotted in Figure 4.

Figure 3. Results of the Placebo tests (non-Annex B donor pool)

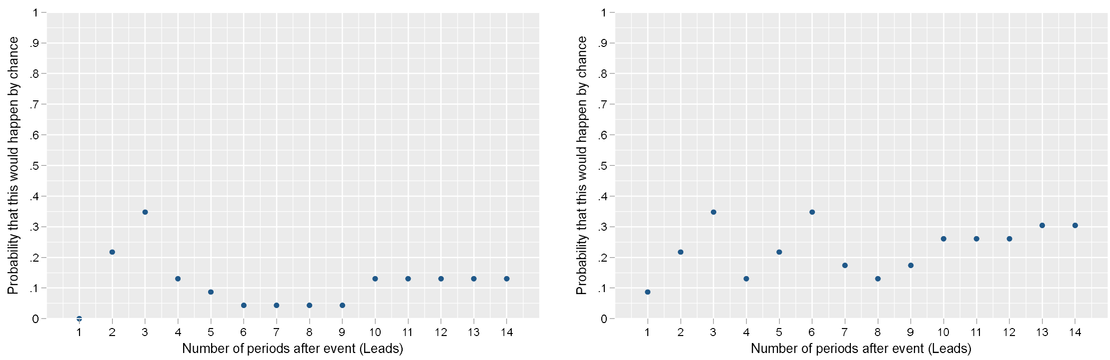
a) My Specification   b) Almer & Winkler Specification



Notes: The solid black line represents the difference between Synthetic New Zealand's CO_2 emissions and the actual CO_2 emissions of New Zealand. The white lines represent the same metric for each of the non-Annex B countries in the donor pool

Figure 4. The p-value plots (non-Annex B donor pool)

a) My Specification   b) Almer & Winkler Specification



Notes: Periods after event (Leads) starts from the treatment year i.e 1=2003. The plotted values represent the pseudo p-value calculated according to equation (10).

The results in Figures 3 and 4 show, that for all post-treatment periods, the estimated treatment effects using the Almer and Winkler (2017) specification are not statistically significant. My specification yields statistically significant treatment effects for the four years ranging from 2008 to 2011. Suggesting that, taking the average treatment effect during this period, legally binding emissions reductions targets resulted in a 21.2% decrease in CO_2 emissions for New Zealand. The ability to produce statistically significant results also suggests that my specification is preferrable to the Almer and Winkler (2017) country level data specification. By extension, this also gives support to the inclusion of CGE modelling variables in the causal analysis of climate policy.

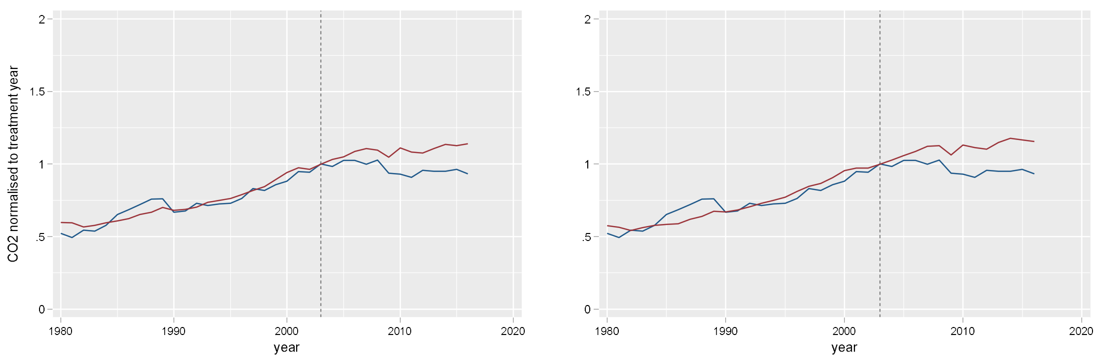
6.2 Control for CDM's

As outlined in the Methodology section, my results must be robust to changes in the design of the study. Furthermore, the Background section outlines the Kyoto Protocol implementation of flexibility mechanisms. The most worrying flexibility mechanism is the CDM, as this directly encourages treated countries to invest in projects which reduce the CO_2 emissions of untreated countries, violating the no interference requirement. To control for the presence of CDMs in the Kyoto Protocol I choose to drop countries from my donor pool which are most likely to have their CO_2 emissions affected by CDMs. China, India, South Korea, Brazil and Mexico are accountable for more than 85% of all certified emission reduction units under the CDM programme (UNFCCC, 2021).[19] I drop the countries that comprise the significant proportion of CDM projects from my donor pool and re-run my specification using my new 'CDM excluded' donor pool.

The results of both my primary specification and the Almer and Winkler (2017) specification, using the CDM excluded donor pool, are presented visually in Figure 5

Figure 5. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B CDM excluded donor pool)

a) My Specification | b) Almer & Winkler Specification



New Zealand
Synthetic Control

RMSPE 0.0472 | RMSPE 0.0557

Notes: The dashed vertical line represents the treatment (ratification in 2003). The RMSPE's for the pre-treatment period are reported under each Figure respectively.

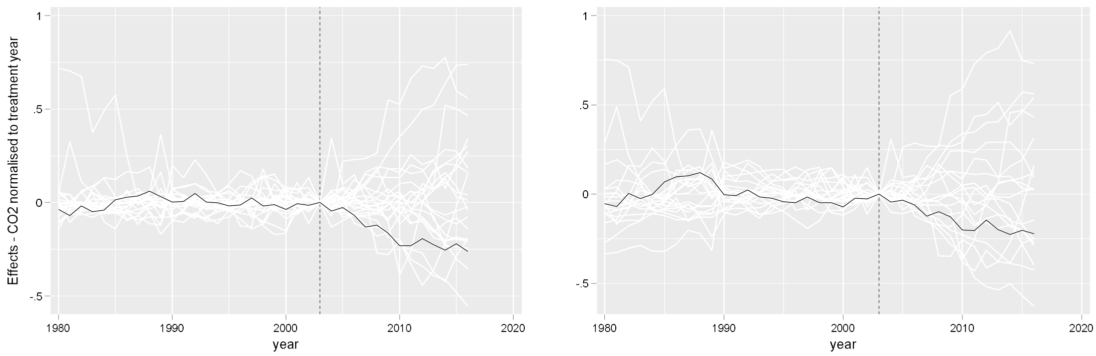
[19] Note that India was never part of the original donor pool due to not meeting the income requirements to be deemed suitable for inclusion.

Using the CDM excluded donor pool, both specifications produce low RMSPEs and track the pre-treatment CO_2 emissions of New Zealand closely. Once again, it is observed that my specification produces a lower RMSPE suggesting that it is preferable in estimating treatment effects of climate change policy. Both specifications estimate similar post-treatment CO_2 emissions paths. My specification suggests that the average treatment effect for the entire post-treatment period is a 10% reduction in CO_2 emissions. The Almer and Winkler (2017) specification suggests that the average treatment effect for the entire post-treatment period is a 12.4% reduction in CO_2 emissions. The large reduction in estimated treatment effects for my specification indicates that it is likely that the excluded donor pool countries had adverse effects in the estimation process.

Looking at the weights used to construct Synthetic New Zealand for both specifications in Figure 2, China is the only country that is given a positive weighting which is now excluded. In my specification China was given a weighting of 0.132 whereas in the Almer and Winkler (2017) specification China only receives a weighting of 0.074. It follows that the large drop in the estimated treatment effect for my specification relative to the Almer and Winkler specification was caused by the large weighting of China biasing the estimated treatment effects. However, this gives no conclusive evidence on the flexibility mechanisms violating the no-interference/SUTVA requirement. In controlling for flexibility mechanisms, primarily CDMs, I would expect an increase in the treatment effect. In theory the CDMs should decrease the GHG, and therefore CO_2, emissions of the CDM host countries, leading to a downward bias in the estimates of the treatment effects. The more probable explanation is the rapid expansion of the CO_2 emissions in China, starting approximately when I assign treatment to New Zealand, results in an upward bias of the treatment effect, proportionate to the weighting given to China. The results of the placebo tests are presented visually in Figure 6 with the results of the subsequent pseudo p-value calculations being plotted in Figure 7.

Figure 6. Results of the Placebo tests (non-Annex B CDM excluded donor pool)

a) My Specification b) Almer & Winkler Specification

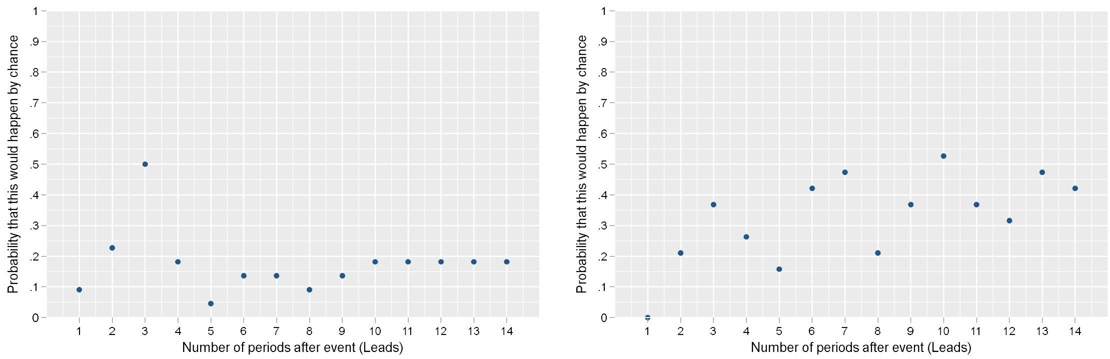


Notes: The solid black line represents the difference between Synthetic New Zealand's CO_2 emissions and the actual CO_2 emissions of New Zealand. The white lines represent the same metric for each of the non-Annex B countries in the donor pool

Figure 7. The p-value plots (non-Annex B CDM excluded donor pool)

a) My Specification

b) Almer & Winkler Specification



Notes: Periods after event (Leads) starts from the treatment year i.e 1=2003. The plotted values represent the pseudo p-value calculated according to equation (10). Values can be interpreted as a standard p-value as per Abadie (2021).

The use of a CDM excluded donor pool appears to reduce the statistical significance of treatment estimates. Once again, the Almer and Winkler (2017) specification fails to produce any statistically significant results. However, my specification also suffers from the reduced donor pool size producing only one statistically significant treatment effect for the year of 2007. Coincidentally the result suggests that legally binding emissions reductions targets produced a statistically significant reduction in CO_2 emissions of 10% for New Zealand in 2007.

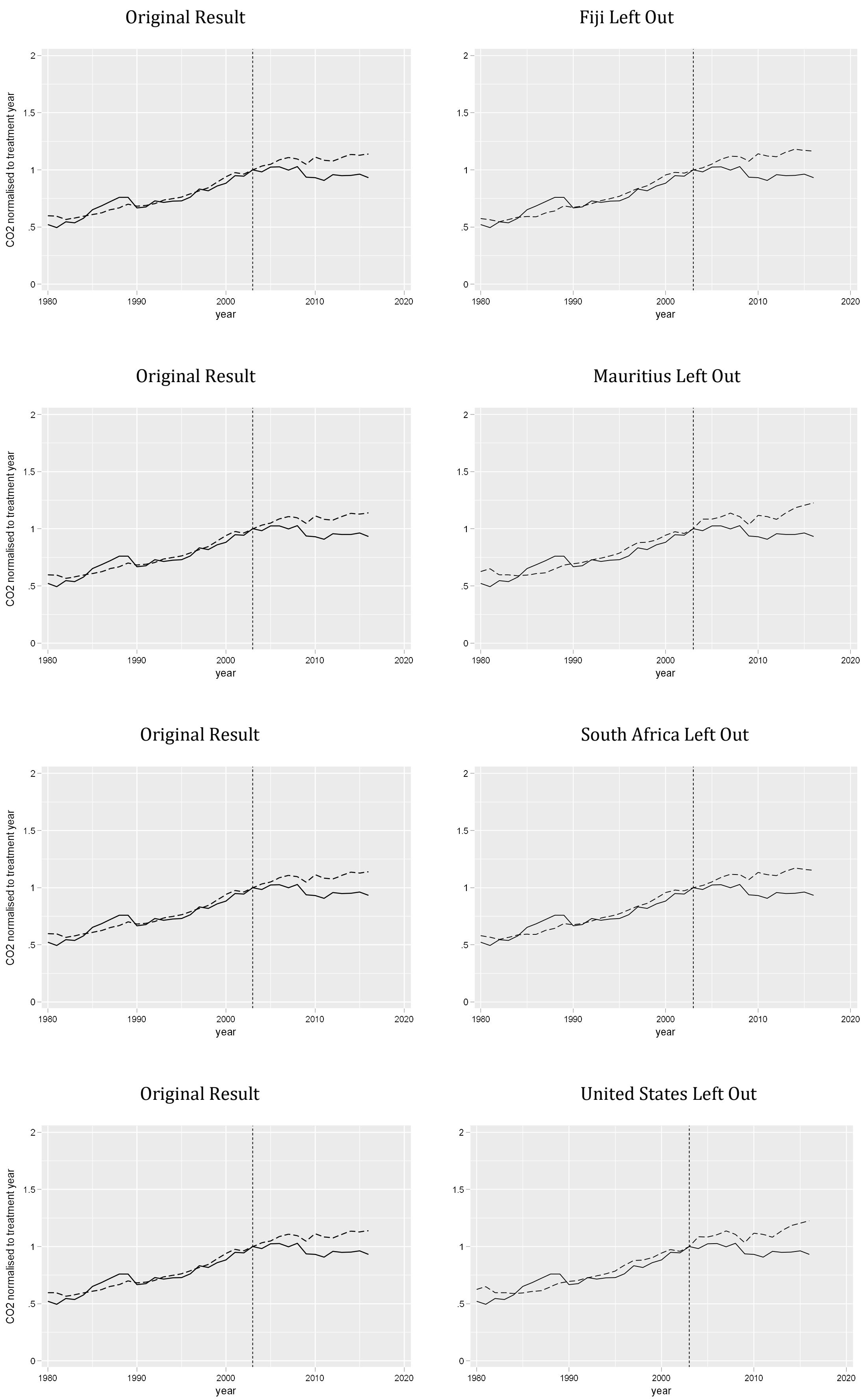
While the lack of statistically significant results is disappointing and highlights the difficulties in estimating causal effects of climate policy, there appears to be consistent evidence of CGE variables improving the quality of the specification.

6.3 Leave-One out tests

To further test the robustness of my results I perform the leave-one tests, based on the weighting prescribed using my specification and the CDM excluded donor pool. Synthetic New Zealand gave positive weighting to four countries: Fiji, Mauritius, South Africa and the US.[20] I iteratively remove each country from the donor pool and re-estimate the Synthetic Control. The results of the leave-one out tests are present visually in Figure 8.

[20] Specifically, Fiji=0.046 Mauritius=0.344 South Africa=0.099 Untied States=0.51

Figure 8. Results of the "Leave-one out" Tests (non-Annex B CDM excluded donor pool)



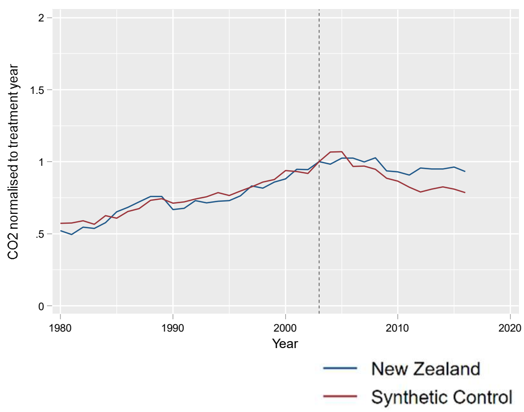
As shown by the leave-one out tests, the results are robust to the removal of any given country with a positive weight. This shows that no one country has an adverse effect on the estimation of the counterfactual outcome showing the results gained so far, limited statistically significant evidence of a treatment effect, are robust.

6.4 US state level data analysis

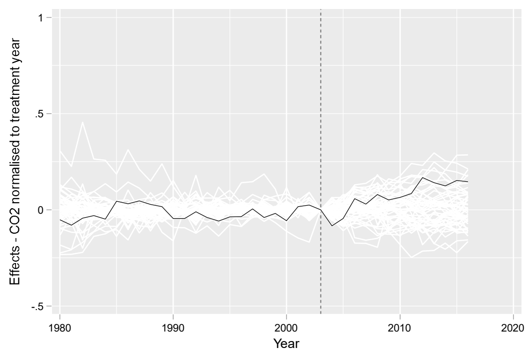
As a further test of the causal impact of Kyoto obligations on the CO_2 emissions of New Zealand I estimate a Synthetic Control using a donor pool comprised of US states. The results of my US state specification are presented visually in Figures 9 and 10.

Figure 9. The Synthetic and Actual CO_2 emissions for New Zealand (US state donor pool)

a) US state specification



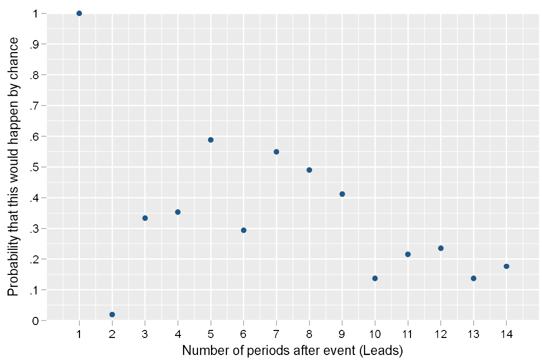
b) placebo tests



RMSPE 0.0498

Notes: The dashed vertical line represents the treatment (ratification in 2003). The RMSPE's for the pre-treatment period are reported under each Figure respectively

c) p-value plot



Notes: Periods after event (Leads) starts from the treatment year i.e 1=2003. The plotted values represent the pseudo p-value calculated according to equation (10). Values can be interpreted as a standard p-value as per Abadie (2021).

When comparing the results for New Zealand using the Almer and Winkler (2017) specifications, the US state donor pool produces a lower RMSPE than the non-Annex B donor pool. This suggests that the US state donor pool is preferable, consistent with the findings in their paper. However, my specification produces a lower RMSPE and lower p-values when using either the complete non-Annex B donor pool or the CDM excluded donor pool.

In contrast to all previous estimates, the US state donor pool suggests that the Kyoto Protocol increased the emissions of New Zealand. While the Almer and Winkler analysis showed that the US state donor pool decreased the effects of the Kyoto Protocol, it was not shown that US state donor pool inverses the treatment effect. To eliminate the potential for coding error, I reproduce the Almer and Winkler (2017) US state donor pool analysis for Australia which is shown in Appendix B. Regardless of the direction of the effect, the US state donor pool also provides no statistically significant evidence of the Kyoto Protocol effecting emissions while also providing evidence for the benefit of including CGE variables in causal estimation of climate policy.

6.5 Summary of Results

The result with strongest evidence is that the inclusion of CGE modelling variables, in a SCM analysis of the effects of the Kyoto Protocol on CO_2 emissions, increases the quality of pre-treatment fit for both New Zealand and Australia.[21] Furthermore, my analysis challenges the findings of Almer and Winkler (2017), which suggest that the use of US state level data is preferable to Country level data. Note that the US state level analysis suggests the Kyoto Protocol had an increasing, albeit statistically insignificant, effect on the CO_2 emissions of New Zealand. Almer and Winkler (2017) had noted that there was a reduction in estimated treatment effects when using US states as the donor pool. While the changed direction of the treatment effect is intriguing, synthetic New Zealand is consistently shown to be more accurately estimated using the country level data. Therefore, I have chosen to focus my discussion on the results obtained from the country level donor pool analysis

Using my extended variable list, I find some evidence of the Kyoto Protocol causing a statistically significant reduction in the CO_2 emissions of New Zealand. However, when I control for the SUTVA violations present due to the CDM there is no statistically significant evidence present. Note that due to the relative recency of the SCM, and the nature of an analysis involving such a small sample size, interpretation of the results requires some nuance. Proper considerations of the results given the issues with the analysis are addressed in the discussion section bellow.

[21] Australia analysis shown in appendix C

Section 7 Discussion

The results of my thesis provide limited statistically significant evidence of the Kyoto Protocol reducing the CO_2 emissions of New Zealand. When I include countries which are heavily involved in CDM projects I find that there is a statistically reduction in the CO_2 emissions of New Zealand for the years 2008, 2009, 2010 and 2011. However, it is preferable to exclude countries heavily involved in CDM projects as their CO_2 emissions are influenced by the Kyoto Protocol and therefore violate the SUTVA condition required for counterfactual analysis. When I use the restricted donor pool, I find no statistically significant evidence of the Kyoto Protocol reducing the CO_2 emissions of New Zealand. The results are consistent with the findings of Almer and Winkler (2017) and the predictions of theoretical literature (Babiker, 2005; Bohringer & Vogt, 2003; Copeland & Taylor, 2005).

The exclusion of countries heavily involved in CDM projects, thus likely to violate SUTVA, provides an unexpected result. In theory countries which host CDM projects should experience lower CO_2 emissions, which is then given to the country which invested in the CDM project as a CER. Therefore, I would expect that CDM countries would experience lower CO_2 emissions than without the Kyoto Protocol resulting in estimated treatment effects which are downwardly biased. It follows that omitting CDM effected countries from the donor pool would increase the estimated treatment effect. Conversely, I find that excluding the countries which account for the majority of CDM projects decreases the treatment effect. Upon further analysis I find that the only CDM effected country which is assigned positive weight is China. The CO_2 emissions of China experienced exponential growth between 1980 and 2016, with the rapid increase beginning around 2003, my assigned treatment date. As the SCM estimates the Synthetic New Zealand using pre-treatment data, weighting is assigned to China during a period where it experienced growth in CO_2 emissions at a similar level to New Zealand. This results in assigning a weighting to China which is relevant for the pre-treatment period, but significantly biases the results in the post-treatment period. The effects seen in China are consistent with the theoretical predictions of significant carbon leakage (Babiker, 2005; Copeland & Taylor, 2005).

In addition, the results of my thesis show that the inclusion of variable used to forecast emissions in CGE literature consistently provides improvements in SCM estimates.[22] To further test this result I conduct the same analysis, comparing specifications, for Australia. The results are given in Appendix C. I show that I can both: reconstruct the results obtained

[22] Specifically, Agricultural land use (% of land area) and Renewable energy consumption (% of total final energy consumption)

from Almer and Winkler (2017); and improve upon their model using my specification. As with the New Zealand analysis, the Australian analysis shows that my specification produces a lower RMSPE and lower average p-value than the Almer and Winkler (2017) specification used in either donor pools. It should be noted that my specification allows the calculation of V weights, and employs a hold-out validation period, whereas Almer and Winkler (2017) use even V weights. However, as shown in Appendix D, the Almer and Winkler specification is largely unchanged when allowing the calculation of V weights as per my specification. This suggests that the primary benefit of allowing the calculation of V weights, and employing a hold-out validation period, is that the specification is tested for robustness in the design phase; eliminating the need for the post estimation in-time placebo tests suggest by Abadie (2021).

It is important to note limitations of my thesis. While the SCM benefits from a data driven process for selecting country weightings, the SCM also requires the selection of variables and pre-treatment time periods of which to match those variables by the researcher. It is my intention to assess both: the effects of Kyoto Protocol on the CO_2 emissions of New Zealand; and the effects of including CGE variables in the causal analysis of climate policy. Therefore, I have chosen to base my analysis off the specification employed by Almer and Winkler (2017). So, although I have established the inclusion of CGE variables is an improvement over a previously employed specification, I have not verified if the previously employed specification was optimal. There is no guidance within the literature as to best practices for selecting predictor variables or time periods of which to average the predictor variables. Future research may look to automate the process of variable and time period selection to reduce the potential for bias in the ad hoc selection process.

Another large caveat to the SCM is the lack of consistency in establishing statistical significance. It is unclear in the literature if the pseudo p-value is an entirely necessary criteria to meet, with the pre-treatment fit often taking hierarchal importance in analysis using the SCM. Even the creator of the method Abadie (2021) advises caution, as the general setting in which the SCM is applicable lends itself to relatively low sample sizes. A small donor pool (< 19) will always estimate treatment effects which are statistically insignificant by the de-facto standard maximum accepted p-value of 0.05. Additionally, the father of null hypothesis Ronald Fisher can be quoted as saying "... no scientific worker has a fixed level of significance at which from year to year, and in all circumstances, he rejects hypotheses; he rather gives his mind to each particular case in the light of his evidence and his ideas."(Fisher, 1973).

When considering the results of my thesis it is important to note that the lack of statistical significance is not conclusive evidence of no effect from the Kyoto Protocol. The results highlight the empirical challenges present in estimating the effects of international environmental policies that can only be evaluated at the country level. Heterogeneity in socioeconomic and political characteristics between countries are difficult to account for simultaneously when employing standard panel data analysis. While the SCM is suited for this type of analysis, the large placebo treatment effects show that the method is not infallible. The interpretation of the results is dependent on how much importance one places on pre-treatment fit and RMSPE ratios. If pre-treatment fit is the primary concern, then we see evidence of the Kyoto Protocol causing a reduction in CO_2 emissions for New Zealand. If RMSPE ratios (pseudo p-values) of < 0.5 are also important, then there is no statistically significant evidence. I stay on the side of caution in wanting to gain both a good pre-treatment fit and acceptable ratio. I stay cautious as the general concern with the SCM is that limited donor pool is what will fail to produce the statistically significant results. However, when I use the US states as a donor pool, I still struggle to obtain statistically significant results.

Further challenges arise when considering the time frame. As noted in literature I reviewed, emissions abatement actions taken today do not always yield immediate results. A longer post treatment may find larger significant treatment effects as the delayed results from abatement efforts begin to show. However, the world is constantly negotiating additional climate agreements, such as the Paris Agreement which entered into force on the 4th of November 2016, which makes establishing suitable comparison units more difficult (*The Paris Agreement | UNFCCC*, n.d.).

Section 8 Conclusion

I estimated the effect of legally binding emissions reductions targets under the Kyoto Protocol on the CO_2 emissions of New Zealand. Employing the SCM and a data set compiled from the World Bank WDI, PWT and Polity IV project, I estimated the CO_2 emissions of New Zealand if it had not been subject to the Kyoto Protocol. In support of the findings of Almer and Winkler (2017), I find no statistically significant evidence that the Kyoto Protocol reduced the CO_2 emissions of New Zealand when compared to a no-Kyoto scenario. The Kyoto Protocol results in business-as-usual emissions as predicted by theoretical literature (Bohringer & Vogt, 2003; Copeland & Taylor, 2005). This conclusion is robust to the omission of positively weighted donor pool countries and an analysis with US state level data. Noteworthy is that the CDMs provide a smaller violation of the SUTVA condition than I assumed. The larger bias comes from the inclusion of China in the donor pool, where CO_2 emissions increased exponentially during the period I analysed (1980-2016), which upwardly biased the estimates.

Additionally, it is shown that the inclusion of variables used to forecast GHG emissions in CGE modelling consistently improves SCM specifications, lowering the RMSPE and average p-value, in all scenarios tested in this thesis. Furthermore, the use of a training and validation period when estimating Synthetic Controls is preferred to forcing equal weighting of predictors as it provides out of sample validation in the design phase while also producing satisfactory predictor imbalances.

Note that the interpretation of the results is dependent on how much importance is placed on obtaining a p-value of < 0.5, an inherently difficult task when undertaking analysis with such small sample sizes. The nature of the analysis, coupled with the lack of consistency around hypothesis testing in the SCM literature leaves some room for interpreting the Kyoto Protocol as having some CO_2 reducing effects. Hopefully future insights provide clarity around best practices in the methodology provide guidelines to interpretating the results of my thesis.

Overall, I do not discredit the difficulty of negotiating an international agreement on climate change. In that regard the Kyoto Protocol is a success and a step in the right direction. However, future climate change agreements should consider the difficulty that the Kyoto Protocol has in producing clear evidence of a reduction in emissions of the participating countries. A more robust design with stronger enforcement mechanisms is necessary if policy makers want to design effective international climate agreements.

References

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

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, *59*(2), 495–510. https://doi.org/10.1111/ajps.12116

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review*, *93*(1), 113–132.

Aldrich, E. L., & Koerner, C. L. (2012). Unveiling Assigned Amount Unit (AAU) Trades: Current Market Impacts and Prospects for the Future. *Atmosphere*, *3*(1). https://doi.org/10.3390/atmos3010229

Almer, C., & Winkler, R. (2017). Analyzing the effectiveness of international environmental policies: The case of the Kyoto Protocol. *Journal of Environmental Economics and Management*, *82*, 125–151. https://doi.org/10.1016/j.jeem.2016.11.003

Andrew, R., & Forgie, V. (2008). A three-perspective view of greenhouse gas emission responsibilities in New Zealand. *Ecological Economics*, *68*(1–2), 194–204. https://doi.org/10.1016/j.ecolecon.2008.02.016

Babiker, M. H. (2005). Climate change policy, market structure, and carbon leakage. *Journal of International Economics*, *65*(2), 421–445. https://doi.org/10.1016/j.jinteco.2004.01.003

Barrett, S. (1998). Political economy of the Kyoto Protocol. *Oxford Review of Economic Policy*, *14*(4), 20–39. https://doi.org/10.1093/oxrep/14.4.20

Beckerman, W. (1972). Environmental policy and the challenge to economic theory. *Social Science Information*, *11*(1), 7–15. https://doi.org/10.1177/053901847201100101

Bennear, L. S., & Coglianese, C. (2005). Measuring Progress: Program Evaluation of Environmental Policies. *Environment: Science and Policy for Sustainable Development*, *47*(2), 22–39. https://doi.org/10.3200/ENVT.47.2.22-39

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Böhringer , C., & Vogt, C. (2003). Economic and environmental impacts of the Kyoto Protocol. *Canadian Journal of Economics/Revue Canadienne D`Economique*, *36*(2), 475–496. https://doi.org/10.1111/1540-5982.t01-1-00010

Böhringer, C., Peterson, S., Rutherford, T. F., Schneider, J., & Winkler, M. (2021). Climate policies after Paris: Pledge, Trade and Recycle: Insights from the 36th Energy Modeling Forum Study (EMF36). Energy Economics, 103, 105471.

Bullock, D. (2012). Emissions trading in New Zealand: Development, challenges and design. *Environmental Politics*, *21*(4), 657–675. https://doi.org/10.1080/09644016.2012.688359

Copeland, B. R., & Taylor, M. S. (2005). Free trade and global warming: A trade theory view of the Kyoto protocol. *Journal of Environmental Economics and Management*, *49*(2), 205–234. https://doi.org/10.1016/j.jeem.2004.04.006

Ertugrul, H. M., Cetin, M., Seker, F., & Dogan, E. (2016). The impact of trade openness on global carbon dioxide emissions: Evidence from the top ten emitters among developing countries. *Ecological Indicators*, *67*, 543–555. https://doi.org/10.1016/j.ecolind.2016.03.027

Faehn, T., Bachner, G., Beach, R., Chateau, J., Fujimori, S., Ghosh, M., Hamdi-Cherif, M., Lanzi, E., Paltsev, S., Vandyck, T., Cunha, B., Garaffa, R., & Steininger, K. (2020). Capturing key energy and emission trends in CGE models: Assessment of Status and Remaining Challenges. *Journal of Global Economic Analysis*, *5*(1), 196–272. https://doi.org/10.21642/JGEA.050106AF

Field, C. B., Vicente R. Barros, Michael D. Mastrandrea, Katharine J. Mach, Monalisa Chatterjee, D.J. Dokken, T.E. Bilir, & M. Chatterjee. (2014). *2014: Summary for Policymakers* (Climate Change 2014: Impacts, Adaptation, and Vulnerability, pp. 1–32). Intergovernmental Panel on Climate Change.

Fisher, R. (1973). *Statistical Methods and Scientific Inference* (3rd ed.).

Gielen, D., & Kram, T. (1998). *The Role of Non-CO2 Greenhouse Gases in Meeting Kyoto Targets*. 13.

Golinski, J. (n.d.). *A Guide to the Climate Change Convention process*. 49.

Grubb, M. (2000). The Kyoto Protocol: An Economic Appraisal. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.229280

Grubb, M. (2004). *Kyoto and the Future of International Climate Change Responses: From Here to Where? 5*(1), 24.

Grunewald, N., & Martinez-Zarzoso, I. (2016). Did the Kyoto Protocol fail? An evaluation of the effect of the Kyoto Protocol on CO $_2$ emissions. *Environment and Development Economics*, *21*(1), 1–22. https://doi.org/10.1017/S1355770X15000091

Gupta, J. (2010). A history of international climate change policy. *WIREs Climate Change*, *1*(5), 636–653. https://doi.org/10.1002/wcc.67

Hartl, A. (2019). The effects of the Kyoto Protocol on the carbon trade balance. *Review of World Economics*, *155*(3), 539–574. https://doi.org/10.1007/s10290-019-00350-5

Imbens, G. W., & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

Intergovernmental Panel on Climate Change. (2014). *Climate Change 2014 Mitigation of Climate Change: Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/CBO9781107415416

Kaul, A., Kloßner, S., Pfeifer, G., & Schieler, M. (2015). *Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates*. 21.

Kim, Y., Tanaka, K., & Matsuoka, S. (2020). *Environmental and economic consequences of the Kyoto Protocol*. 36.

Leining, C., & Kerr, S. (2018). *A guide to the New Zealand Emissions Trading Scheme* (p. 15). Motu Economic and Public Policy Research.

Leining, C., Kerr, S., & Bruce-Brand, B. (2020). The New Zealand Emissions Trading Scheme: Critical review and future outlook for three design innovations. *Climate Policy*, *20*(2), 246–264. https://doi.org/10.1080/14693062.2019.1699773

Levinson, A. (1995). *Environmental Regulation and Industry Location: International and Domestic Evidence*. Department of Economics University of Wisconsin. https://www.researchgate.net/publication/246078202
Maamoun, N. (2019). The Kyoto protocol: Empirical evidence of a hidden success. *Journal of Environmental Economics and Management*, *95*, 227–256. https://doi.org/10.1016/j.jeem.2019.04.001
Means, T., & Lallanilla, M. (2021, June 9). *Greenhouse gases: Causes, sources and environmental effects*. Livescience.Com. https://www.livescience.com/37821-greenhouse-gases.html
Molina, M., & McCarthy, J. (2014). *THE AAAS CLIMATE SCIENCE PANEL*. 14.
New Zealand & Interim Climate Change Committee. (2019). *Action on agricultural emissions: Evidence, analysis and recommendations, 30 April 2019*. https://www.iccc.mfe.govt.nz/what-we-do/agriculture/agriculture-inquiry-final-report/
New Zealand Ministry for the Environment. (2015, April 1). *New Zealand's Greenhouse Gas Inventory 1990–2013*. Ministry for the Environment. https://environment.govt.nz/publications/new-zealands-greenhouse-gas-inventory-19902013/
New Zealand Ministry for the Environment. (2021, May 4). *New Zealand and the United Nations Framework Convention on Climate Change*. Ministry for the Environment. https://environment.govt.nz/what-government-is-doing/international-action/nz-united-nations-framework-convention-climate-change/
Nielsen, T., Baumert, N., Kander, A., Jiborn, M., & Kulionis, V. (2021). The risk of carbon leakage in global climate agreements. *International Environmental Agreements: Politics, Law and Economics*, *21*(2), 147–163. https://doi.org/10.1007/s10784-020-09507-2
Nunez, C. (2019, January 22). *What Is Global Warming?* National Geographic. https://www.nationalgeographic.com/environment/article/global-warming-overview
Polity IV Project: Home Page. (n.d.). Retrieved 2 August 2021, from http://www.systemicpeace.org/polity/polity4x.htm
PWT 10.0. (2021, January 7). University of Groningen. https://www.rug.nl/ggdc/productivity/pwt/
Richter, J. L., & Chambers, L. (2014). Reflections and outlook for the New Zealand ETS: Must uncertain times mean uncertain measures? *Policy Quarterly*, *10*(2). https://doi.org/10.26686/pq.v10i2.4485
Ritchie, H., & Roser, M. (2020). CO_2 and Greenhouse Gas Emissions. *Our World in Data*. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions
Rose, S. K., & Sohngen, B. (2011). Global forest carbon sequestration and climate policy design. *Environment and Development Economics*, *16*(4), 429–454.
Rubin, D. B. (1980). Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment. *Journal of the American Statistical Association*, *75*(371), 591–593. https://doi.org/10.2307/2287653
Shahbaz, M., & Sinha, A. (2019). Environmental Kuznets curve for CO2 emissions: A literature survey. *Journal of Economic Studies*, *46*(1), 106–168. https://doi.org/10.1108/JES-09-2017-0249
Stats NZ. (2020, October 15). *New Zealand's greenhouse gas emissions | Stats NZ*. Stats NZ. https://www.stats.govt.nz/indicators/new-zealands-greenhouse-gas-emissions

Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development*, *32*(8), 1419–1439. https://doi.org/10.1016/j.worlddev.2004.03.004

The Economist. (2019, September 19). The climate issue. *The Economist*. https://www.economist.com/leaders/2019/09/19/the-climate-issue

The Paris Agreement | UNFCCC. (n.d.). Retrieved 13 August 2021, from https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement

UNFCCC. (2008). *Kyoto Protocol Reference Manual: On accounting of emissions and assigned amount* (p. 130) [Reference Manual]. UNFCCC. https://unfccc.int/sites/default/files/08_unfccc_kp_ref_manual.pdf

UNFCCC. (2021, June 30). *CDM insights*. https://cdm.unfccc.int/Statistics/Public/CDMinsights/index.html

United Nations. (1992). *The United Nations Framework Convention on Climate Change* (p. 33).

United Nations. (1998). *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. https://unfccc.int/resource/docs/convkp/kpeng.pdf

U.S. Bureau of Economic Analysis. (2021). *GDP by State | U.S. Bureau of Economic Analysis (BEA)*. https://www.bea.gov/data/gdp/gdp-state

U.S. Census Bureau. (2021). *Population and Housing Unit Estimates*. The United States Census Bureau. https://www.census.gov/popest

U.S. Energy Information Administration. (2021). *State Carbon Dioxide Emissions Data*. https://www.eia.gov/environment/emissions/state/index.php

U.S. Environmental Protection Agency. (2015, December 23). *Overview of Greenhouse Gases* [Overviews and Factsheets]. https://www.epa.gov/ghgemissions/overview-greenhouse-gases

Wing, I. S. (2009). *Computable general equilibrium models for the analysis of energy and climate policies. International handbook on the economics of energy*. World Bank. (2020). *World Development Indicators | DataBank*. https://databank.worldbank.org/source/world-development-indicators

World Bank. (2021). *WDI - Sources and Methods*. https://datatopics.worldbank.org/world-development-indicators/sources-and-methods.html

York, R. (2005). Kyoto Protocol Participation: A Demographic Explanation. *Population Research and Policy Review*, *24*(5), 513–526. https://doi.org/10.1007/s11113-005-1291-x

Zahran, S., Kim, E., Chen, X., & Lubell, M. (2007). Ecological Development and Global Climate Change: A Cross-National Study of Kyoto Protocol Ratification. *Society & Natural Resources*, *20*(1), 37–55. https://doi.org/10.1080/08941920600981355

Zhang, J., & Wang, C. (2011). Co-benefits and additionality of the clean development mechanism: An empirical analysis. *Journal of Environmental Economics and Management*, *62*(2), 140–154. https://doi.org/10.1016/j.jeem.2011.03.003

Appendices

Appendix A. Treatment effects by year

Table A1. Treatment effects by year

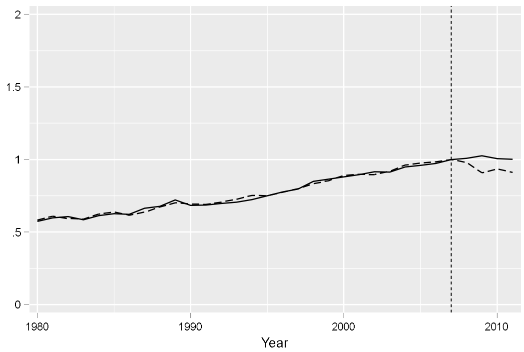
Year	My Specification Treatment effect	Almer and Winkler Treatment effect
Non-Annex B Countries		
Entire Donor Pool		
2004	-0.04	-0.05
2005	-0.05	-0.04
2006	-0.09	-0.07
2007	-0.15	-0.14
2008	-0.12 *	-0.11
2009	-0.18 *	-0.15
2010	-0.26 *	-0.23
2011	-0.29 *	-0.24
2012	-0.23	-0.18
2013	-0.28	-0.23
2014	-0.29	-0.25
2015	-0.27	-0.23
2016	-0.29	-0.25
RMSPE	0.042	0.052
CDM Control		
2004	-0.04	-0.04
2005	-0.02	-0.03
2006	-0.06	-0.06
2007	-0.10 *	-0.12
2008	-0.06	-0.09
2009	-0.11	-0.12
2010	-0.18	-0.20
2011	-0.17	-0.20
2012	-0.11	-0.14
2013	-0.15	-0.19
2014	-0.18	-0.22
2015	-0.16	-0.20
2016	-0.20	-0.22
RMSPE	0.047	0.056

Notes: Asterisks indicate the significance level of the estimated treatment effect where ** if p<0.01 and * if p<0.05
Source: My own calculations based on data from WDI, PWT and Polity IV project.

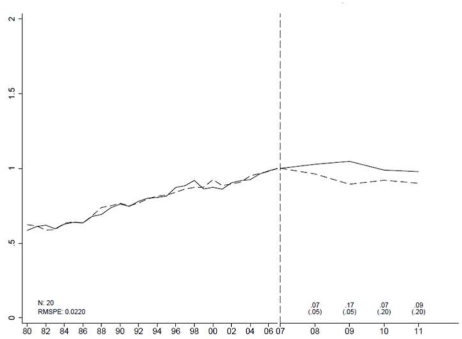
Appendix B. Recreating the Almer and Winkler (2017) result of Australia

Figure B1. The Synthetic and Actual CO_2 emissions for Australia (US State donor pool)

a) My Result



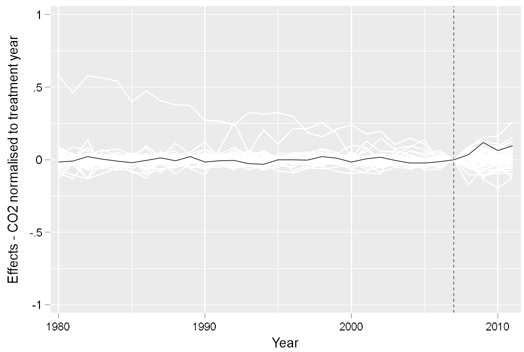
b) Almer and Winkler's Result



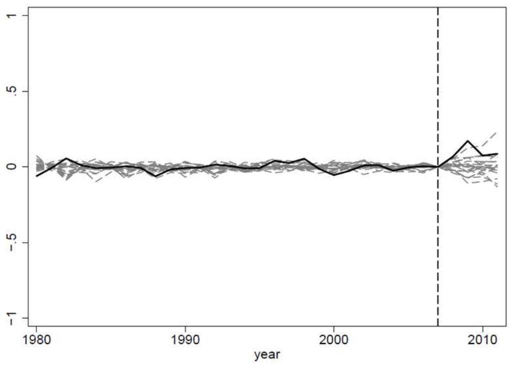
Notes: In both graphs that solid black Visual differences arise due to the X-axis scaling. My results use a even scale whereas Almer and Winkler choose to compress the scale of the pre-Kyoto outcomes and extend the scale of the post-Kyoto outcomes.

Figure B2. Results of the Placebo tests (US State donor pool)

a) My Specification



b) Almer & Winkler Specification



Notes: In both graphs the solid black line represents the difference between Synthetic Australia's CO_2 emissions and the actual CO_2 emissions of Australia. In graph a) the white lines represent the same metric for each of the US states in the donor pool. Likewise, in graph b) the dashed grey lines represent the same metric for each of the US states in the donor pool. The presence of extra placebos in my graph is explained by Almer and Winkler option to only graph the results of placebo tests of the 19 States with the lowest pre-Kyoto RMSPE.

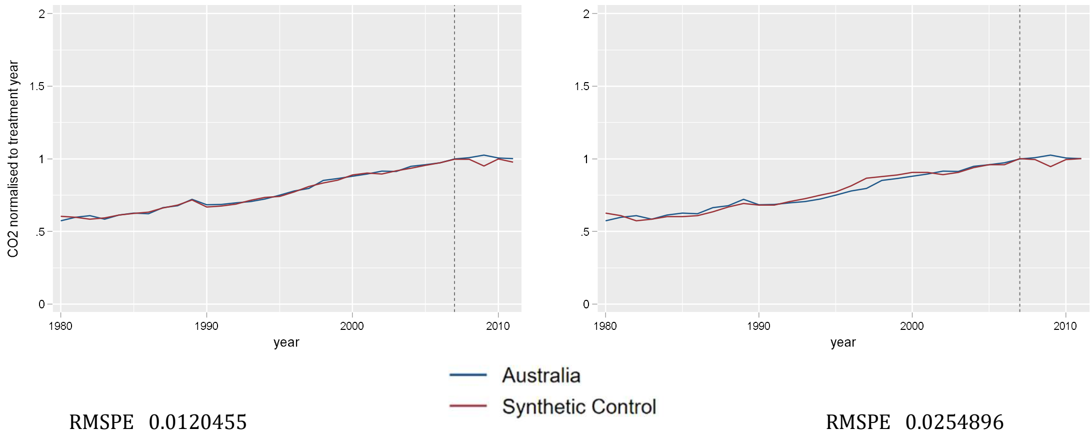
Appendix C. Comparing the specifications using Australia as the treated country

To establish that the improvement of specification quality due to inclusion of CGE variables is not specific to New Zealand, I reconstruct the Almer and Winkler (2017) results for Australia while also applying my specification. To increase comparability, I restrict the year range to match their study exactly (1980-2011). The results of both my primary specification and the Almer and Winkler specification are presented visually in Figure C1

Figure C1. The Synthetic and Actual CO_2 emissions for Australia (non-Annex B donor pool)

a) My Specification

b) Almer & Winkler Specification



RMSPE 0.0120455

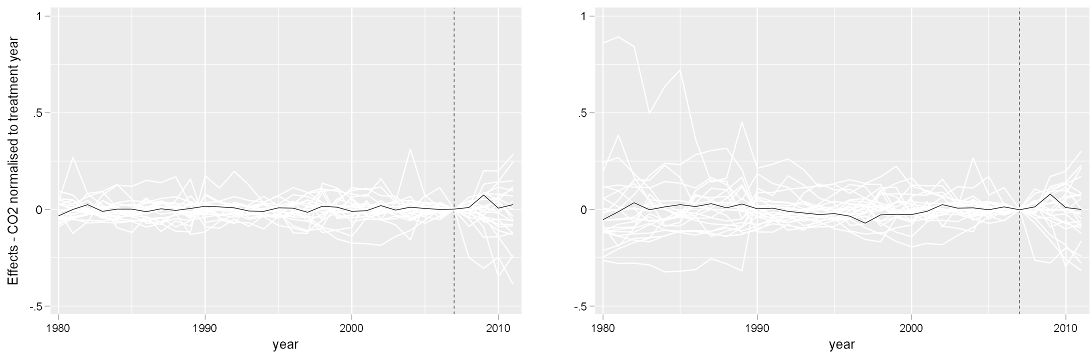
RMSPE 0.0254896

The results of the placebo tests are presented visually in Figure C2 with the results of the subsequent pseudo p-value calculations being plotted in Figure C3.

Figure C2. Results of the Placebo tests (non-Annex B donor pool)

a) My Specification

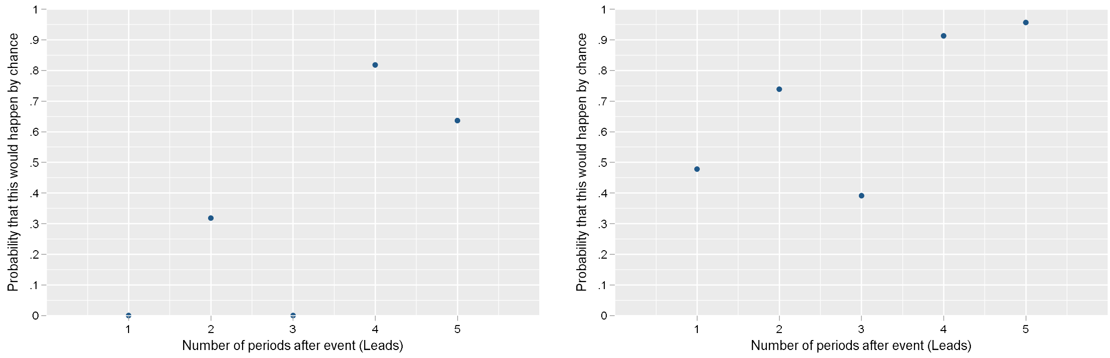
b) Almer & Winkler Specification



Notes: The solid black line represents the difference between Synthetic Australia's CO_2 emissions and the actual CO_2 emissions of Australia. The white lines represent the same metric for each of the non-Annex B countries in the donor pool

Figure C3. The p-Value plots (non-Annex B donor pool)

a) My Specification b) Almer & Winkler Specification



Notes: Periods after event (Leads) starts from the treatment year i.e 1=2007. The plotted values represent the pseudo p-value calculated according to equation (10). Values can be interpreted as a standard p-value as per Abadie (2021).

Unfortunately, I have been unable to provide any robust statistically significant evidence for a reduction in CO_2 emissions attributable to the Kyoto Protocol using either specification on country level data. However, my specification consistently produces lower RMSPE for the pre-treatment period and exhibits lower probability of the estimated treatment effects happening by chance. Adding to the evidence that the inclusion of CGE variables in climate policy causal analysis is worthwhile.

Appendix D. Comparison of V weight methods

Below are comparisons of the Almer and Winkler (2017) non-Annex B donor pool analysis using both: calculated V weights; and even V weights.

Figure D1. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B donor pool)

a) Almer & Winkler specification　　　　b) Almer & Winkler Specification (calculated V weights)

Figure D2. The Synthetic and Actual CO_2 emissions for New Zealand (non-Annex B CDM excluded donor pool)

a) Almer & Winkler specification　　　　b) Almer & Winkler Specification (calculated V weights)

As shown by Figures D1 and D2, the introduction of calculated V weights produce near identical outputs. This suggests that the vast majority of difference between my specification and Almer and Winkler (2017) is accounted for by the introduction of the CGE variables as opposed to the choice of V weights.