Measuring the Impact of Cap-and-Trade Policies on Carbon Emissions Reduction in the United States: A Quantitative Analysis Design

Jiawei Sun, Xinyu Zheng, Ming Zhou

1. Introduction

1.1 Backgrounds

It is established by the United Nations that we are experiencing a climate emergency that has been threatening the lives, economy, health, and food of humanity. The concentration of greenhouse gases (GHG) in the atmosphere has significantly altered the composition of the atmosphere and has heated the average world temperature. Despite the consensus of reducing GHG emissions, the gap between aspiration and reality is far from being achieved. The imperativeness and exigency of interventions become even more important if we were to realize the goal set forth by the Paris Agreement.

Multiple policy instruments have been developed to achieve the goal of reducing emissions. Unlike "command-and-control" approaches where governments use regulations to standardize or promulgate technology choices for individual facilities and businesses. Cap-and-Trade (CAT) program allows the market to set a price on carbon and that price drives investment decisions and spurs market innovations.

In a CAT system, the government would set an emission cap and issue a quantity of total emission allowances along with that cap. Emitters must hold allowances for every ton of greenhouse gas they emit. Allowances can be bought and sold for an emission price determined by the market. Companies, therefore, have been incentivized to reduce emissions and control the cost by selling any excess allowances. Meanwhile, companies will face higher costs to buy the allowances if they fail to curb their emissions.

It's worth mentioning that CAT is the preferred policy tool for a carbon tax when the government has a specified emission target. The distinction between CAT and a carbon tax lies in the level of predictability they provide with regard to emissions. Carbon taxes provide a definite cost for emissions; however, they do not provide a clear estimation of the levels of emissions that would result from that cost. Conversely, CAT programs establish a definite projection of future emissions, they do not establish a definite price for these emissions.

1.2 Motivation

Being one of the world's highest per capita GHG emissions countries, the United States has started using market-driven policy tools like CAT schemes to incentivize GHG emission reduction. Twelve states that are home to over a quarter of the U.S. population and account for a third of the U.S. GDP have instituted an active carbon-pricing program, the Regional Greenhouse Gas Initiative (RGGI). Being the first mandatory CAT program, the program limits the carbon dioxide emissions from the power sector in the participating states. In 2013, California started to operate a CAT program that became the first multi-sector CAT program in North America. The program has been established and in action for more than a decade (RGGI, 2019).

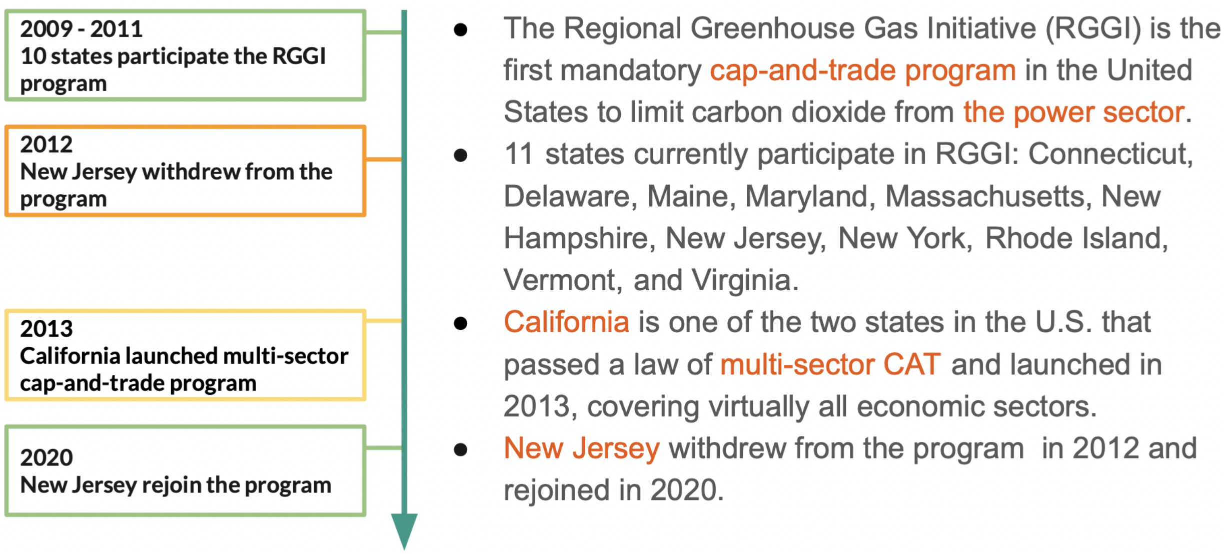

Figure 1. Timeline and introduction of Cap-and-Trade program in U.S.

Policy briefs and media commentary have generally portrayed this CAT program in a well-regarded manner. However, we must answer for the effectiveness and the impacts of the program before we push forward the CAT program as a nationwide policy. Thus in this proposal, we seek to evaluate the effectiveness of this ongoing CAT program in reducing carbon emissions.

1.3 Problem Statements

In our research, we aim to evaluate the effect of CAT schemes on reducing carbon emissions. We thus have three research questions:

(1) Is the multi-sector CAT scheme in California more effective in reducing carbon emissions compared to states that only applied CAT in the power sector?

(2) Does the power-sector CAT have a crowd-out effect? In other words, are there differences in emissions in the power and non-power sectors in the RGGI states before and after the CAT?

(3) Does CAT have a cross-border effect? In other words, will CAT in one state affect the emission in neighboring states?

2. Conceptual Framework

2.1 Market Failure

A market failure refers to the inefficient distribution of resources that occurs when the individuals in a group end up worse off than if they had not acted in rational self-interest. For decades, many economists have taken climate change as a classical example of market failure. The core one is the so-called 'greenhouse-gas externality'. It indicates that greenhouse gas emissions are an external effect of a range of economically valuable activities, including burning coal to generate electricity, burning petrol to power cars, producing food and disposing of waste. Most of the impacts of emissions would not fall on those conducting the activities, rather, they would fall on future generations or people living in low-income or middle-income countries (Ehigiamusoe & Lean, 2019). In other words, those who are responsible for the emissions do not pay the cost. The adverse effects of greenhouse gasses are therefore negative externalities to the whole society. Since there is no incentive for businesses and consumers to reduce emissions, the market fails by over-producing greenhouse gasses. Our intervention aims to create incentives in the market (on businesses mainly) to produce less greenhouse gasses, generating less deadweight loss and reaching market equilibrium.

2.2 Theory of Change

Currently, the major greenhouse gas reduction policy approaches fall into three main categories: carbon pricing, technology subsidies, and performance standards. We choose multi-sector/single-sector CAT programs based on political feasibility and its long-term economic benefits. A CAT program is a common government regulatory program designed to limit or cap the total amount of emissions of carbon dioxide, as a result of industrial activity. It is a market system since it creates an exchange value for emissions. In general, the government would set the limit on emissions permitted across industries. It would issue a limited number of permits that allow companies to emit a certain amount of carbon dioxide. And the total amount of the cap will be split into allowances. Each allowance permits a company to emit one ton of emission. If a company wants to emit more greenhouse gasses than they are allotted, they must buy allowance credits from the state during an auction. The proceeds from these sanctions go toward other climate projects, such as land restoration and conservation. In this way, companies would be ‘taxed’ if they produce a higher level of emissions than their permits allow. The government may also lower permits each year, making it more expensive to emit. A CAT program offers an incentive for companies to reduce emissions by investing in cleaner technologies or funding research into alternative energy resources. It can also lead to faster cuts in emission.

2.3 Relevance

The local impacts of a CAT program can provide evidence in its effectiveness in reducing emissions in either long term or short term, and its effectiveness relative to the program in single sector (e.g., limiting to the energy sector), empirical knowledge in setting the limits and caps, and the speed of lowering them. There are also other lessons for other institutions:

- How substantial the revenue generated by auctioning its allowances will be， especially for the countries/states, where their governments are in budget deficits.
- Empirical lessons in setting the ‘price’ of allowances in firms across sectors.
- How a CAT program in carbon dioxide can be jointly implemented with other programs in local economies.

3. Research Design

3.1 Intervention and sample

In total, thirteen states in the U.S. adopted CAT programs. California is one of the two states in the U.S. that has passed a law of multi-sector CAT and launched in 2013, covering virtually all economic sectors. Washington is the other state that implemented a multi-sector CAT program, but it was just taking effect in 2023. As of today, eleven states have launched the CAT programs under the RGGI since 2009, targeting the power sector only. New Jersey dropped out of RGGI in 2012 and rejoined in 2020.

Stemming from the context of the CAT program, the interventions for the three research questions are listed as follows:

(1) To evaluate whether a multi-sector CAT program is more effective than a single-sector one, the intervention can only be the California multi-sector CAT program.
(2) To test the crowd-out effect of single-sector CAT on other sectors, the intervention can only be the RGGI power-sector CAT programs.
(3) To investigate any cross-border effect of CAT programs, the intervention can be either the California multi-sector CAT program or RGGI power-sector CAT programs.

The effectiveness of CAT programs, the crowd-out effect, and the cross-border effect are all measured by the changes in GHG emissions due to the corresponding interventions.

3.2 Data

As explained in 3.1, the efficiency of multi-sector CAP programs, the crowd-out effect, and the cross-border effect are measured by comparing the reduction in GHG emissions by facilities across intervention status, sectors, and geographic locations. Thus, we plan to use observational data from the greenhouse gas reporting program (GHGRP).

The GHGRP requires reporting greenhouse gas (GHG) data in the United States. Approximately 8,000 facilities have been required to report their emissions annually since 2010. It covers a total of 41 categories of reporters. The Environmental Protection Agency (EPA) determines which facilities to submit annual emission reports. Facilities that:

(1) GHG emissions from covered sources exceed 25,000 metric tons of CO_2 per year.
(2) The supply of certain products would result in over 25,000 metric tons of CO_2 in GHG emissions if those products were released, combusted, or oxidized.

Facilities are defined as stationary sources of greenhouse gas emissions. Facilities covered by the GHGRP include 9 sectors, such as power plants, oil and gas production and refining facilities, manufacturing plants, waste management facilities, and others. The GHGRP covers emissions from both direct sources (such as emissions from a power plant's smokestack) and indirect sources (such as emissions from the electricity consumed by a manufacturing plant). Since one company may have various facilities within one state or have multiple facilities across states, we will aggregate facility-level emission data into company-level emission data. A company-state's GHG emission is calculated as the total GHG emissions of all the facilities owned by the company within one state. If the company has facilities in multiple states, we will generate company-level total emissions data for each state.

Our sample is drawn from facilities in California, RGGI states, and neighboring states, including Arizona, Nevada, Oregon.

The GHGRP specifies methodologies for calculating GHG emissions from each source category. Reporters can generally choose from a number of methods for calculating GHG emissions. Existing environmental monitoring systems and other factors may influence the choice of method. The EPA will conduct electronic validation and verification checks on all reports. The EPA will notify the reporter if any potential errors are discovered.

Facility-level GHG data covers facility name (and ID), location (State, city, and zip code), NAICS codes (Industrial Classification code), GHGRP sector (including 14 industries), and greenhouse emissions (including CO_2, CH_4, N_2O, HFC, PFC, SF_6, NF_3, HFE, and others). The following table shows the description of facility information and emissions data.

Table 1: The documentation of the Greenhouse Gas Reporting Program dataset

Variable	*Description*	*Metric*
Facility Data		
Facility Name	The name of the Facility	Character
Facility ID	Unique GHGRP identifier for facility	Numeric
Location_state	Which state the facility located in	Character
Location_city	Which city the facility located in	Character
Location_county	Which county the facility located in	Character
Location_zip_code	The zip codes of the facility	Numeric
Sector	Power plants; Refineries; Non-Fluorinated Chemicals; Minerals; Petroleum and Natural gas; Waste; Pulp and Paper; Miscellaneous combustion; Metals; Fluorinated Chemicals; Underground Coal Mines; Electrical Equipment; Electronics Manufacturing	Character
Emission Data		
Total_Emission	Total greenhouse gas Emission in millions of tons	Numeric
CO2	CO2 Emission in millions of tons	Numeric
CH4	CH4 Emission in millions of tons	Numeric
N2O	N2O Emission in millions of tons	Numeric
HFC	HFC Emission in millions of tons	Numeric
PFC	PFC Emission in millions of tons	Numeric
SF6	SF6 Emission in millions of tons	Numeric
NF3	NF3 Emission in millions of tons	Numeric
HFE	HFE Emission in millions of tons	Numeric
F_GHG	Other fully fluorinated GHG Emission in millions of tons	Numeric
Lived_compounds	Very short lived compounds Emission in millions of tons	Numeric

3.3 Recall Survey

While the GHGRP provides useful emissions data, it only applies to large facilities emitting more than 25,000 metric tons of CO_2. Because GHGRP excludes small facilities, our understanding of the overall emission landscape is limited. Besides, there was no additional data on emission-related information such as revenue, energy efficiency, R&D investment, and company use of high technology.

To fill this gap, we would create a recall survey that would collect emission data from small facilities as well as additional relevant information from all facilities. Collecting emission data from small facilities is critical since they account for a significant portion of greenhouse gas emissions. Moreover, understanding their emission patterns and drivers is essential to designing effective mitigation strategies. By incorporating this additional data into our models, we would be able to control for time-various company characteristics and gain an unbiased understanding of greenhouse gas emissions across all facility sizes.

3.3.1 Survey Questions

Our sample states includes: (1)RGGI states: Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey (withdrew in 2012, rejoined in 2020), New York, Rhode Island, Vermont, and Virginia, (2)California, (3)Non-RGGI states that are adjacent to California: Arizona, Nevada and Oregon.

For all the companies in the sample states, we will collect companies' basic information, revenue, energy efficiency, R&D investment, and companies' use of high technology prior to CAT implementation. For all the sample states except New Jersey, we will collect emission data of the small facilities from 2012 to 2016. For New Jersey specifically, we will collect emission data of the small facilities from 2019 to 2020.

Table 2: Summary of recall survey questionnaire

Designed Questions	
Basic Information	
1. The ID of the company.	
2. Which state the company locates?	
3. The Zip code of the company location.	
4. Which sector the company is in?	
5. How many facilities you would like to report?	
6. Please provide each facility ID and the state it locates.	
Emission	
7. The ID of the facility.	
8*. The GHG emission between 2010 – 2016 (if not in New Jersey) in metric tons.	
9*. The GHG emission between 2019– 2021 (if in New Jersey) in metric tons.	
Other Relevant Information	
10. The ID of the company.	
11. Are the reported facilities in the same state?	
12. If not, please provide each state it locates and answer the below question separately for each state.	
Revenue	13*. The revenue between 2010 – 2016 (if not in New Jersey) in millions of U.S. dollars.
	14*. The revenue between 2019 – 2021 (if in New Jersey) in millions of U.S. dollars.
Energy Efficiency	15*. The energy efficiency between 2010 – 2016 (if not in New Jersey) in %.
	16*. The energy efficiency between 2019 – 2021 (if in New Jersey) in %.
Research & Design Investment	17*. The research and design investment between 2010 – 2016 (if not in New Jersey) in millions of U.S. dollars.
	18*. The research and design investment between 2019 – 2021 (if in New Jersey) in millions of U.S. dollars.
	19*. Whether the company implemented new technology before the CAT implementation.

Notes: To generate a comprehensive data set for our recall survey, we use the 'repeat group' design in SurveyCTO. This design enables us to generate a repeat group for each year, facility, and state that respondents report on and to capture data on greenhouse gas emissions as well as other relevant information. This design will be implemented in the questions marked with an asterisk (*) to ensure that we collect data for each relevant category.

3.3.2 Survey Process

The graph below depicts the survey collection process. To conduct a recall survey, we will first identify and collect contact information from companies registered in the GHGRP. We will contact these companies and disseminate our survey through email. A reminder email will also be sent 20 business days before the survey ends. After survey responses have been gathered, they will be cleaned and used for additional data analysis.

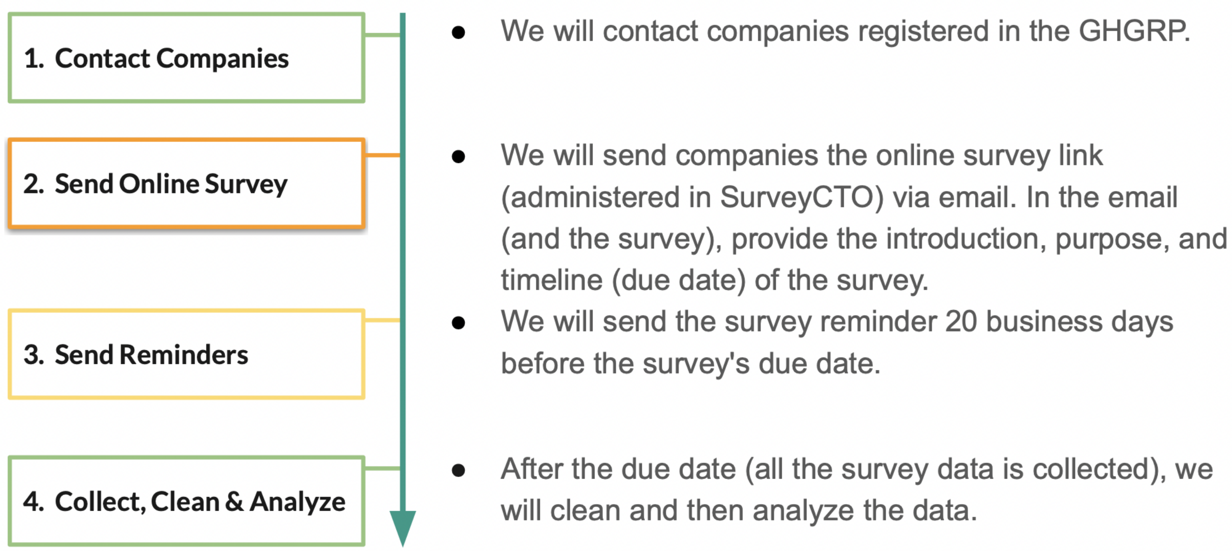

Figure 2: Recall survey process

3.3.3 Survey Data Validation

To ensure the accuracy and quality of our survey data, we will perform a range of validation checks. This includes looking for missing values and outliers, as well as verifying survey duration times and looking for duplicates.

- The missing values and outliers check will help us to identify and address any instances where respondents left questions unanswered, ensuring that our data is as complete as possible. The outlier check will detect and flag any responses that are outside of the expected ranges, allowing us to review and potentially correct any errors or anomalies. In SurveyCTO, we will also set up the following checks.
    - Set range checks for all numeric data to ensure that it falls within a given range.
    - Set up Zip code validation expressions to ensure that the data entered meets certain Zip code criteria.
    - Set up required questions to ensure that all emission questions are answered

before respondents submit the survey.
    - Set up dynamic checks to ensure that the number of the repeated groups matches the previous questions, such as the number of facilities the respondent has reported and the state where the facility is located.
- Additionally, we will validate survey duration times to ensure that responses are not submitted too quickly or too slowly, which may result in data quality issues.
- Finally, we will use the email id to identify and remove any instances where respondents may have submitted multiple surveys.

Through these validation checks, we can ensure that our survey data is reliable and of high quality, allowing us to draw meaningful insights and conclusions.

3.4 Methodology

To answer the research questions proposed in section 1, we employ a Difference-in-Difference framework: (1) Compare the GHG emissions in the state(s) implemented multi-sector CAT programs (treatment group) with those that adopted power-sector programs (control group) before and after the program implementation; (2) In state(s) implemented power-sector CAT programs, compare the GHG emissions in the state(s) in the power sector (treatment group) with those in the non-power sector (control group) before and after the program implementation; (3) Compare the GHG emission in the state(s) neighboring on the CAT adopter(s) (treatment group) with those in neighboring states that have not adopted any CAT programs (control group) before and after the program implementation.

Function form is written as follow (Bartram et al., 2022):

$$Emission_{c,y} = \alpha + \beta Treatment_c \times Post_y + \gamma' X_{c,y} + a_c + b_y + \varepsilon_{c,y}$$

Where $\boldsymbol{Emission_{c,y}}$ is the metric tons of CO2 emitted by $\boldsymbol{Company_c}$ in $\boldsymbol{Year_y}$; $\boldsymbol{Treatment_c}$ is an indicator variable equal to 1 if the company is in the treatment group, and 0 otherwise; $\boldsymbol{Post_y}$ is an indicator equal to 1 if the year is after the implementation of the Cap-and-Trade policy (2013); $\boldsymbol{X_{c,y}}$ indicates a set of control variables, including revenue, energy efficiency, and Research & design investment; $\boldsymbol{a_c}$ and $\boldsymbol{b_y}$ each denotes company fixed effects and year fixed effects. The variables $\boldsymbol{Treatment_c}$ and $\boldsymbol{Company_c}$ are not included by themselves in the regressions, as they are subsumed by the fixed effects. We adjust standard errors for clustering at the company and year levels. To study the impact of Cap-and-trade policy on the emission of greenhouse gas, we evaluate whether the coefficients on the interaction term are statistically significant at the level of 0.05.

Based on GHGRP data availability, research questions, and corresponding interventions, we choose samples of treatment and control groups as follows:

(1) To analyze the validity of a multi-sector CAT program versus a single-sector CAT program, California facilities covered in GHGRP are the only choices of the treatment group, and the facilities covered in GHGRP in all RGGI states except New Jersey are included in the control group. As the California CAT program was launched in 2013, we take the data from 2010, three years before the treatment, to 2016, three years after the treatment.

Table 3: Difference-in-Difference Model for Research Question 1

	Pre treatment	Post treatment
Treatment: California companies	2010 - 2013	2013 - 2016
Controlled: companies in other 9 states	2010 - 2013	2013 - 2016

(2) To see whether there is a crowd-out effect across sectors, considering the GHGRP data before 2009 when the power-sector CAT under RGGI was implemented, are unavailable, we can take advantage of the New Jersey state, which dropped out the RGGI in the middle and rejoined in 2020. We take facilities covered by GHGRP in the power sector of New Jersey as the treatment group, while those in the non-power sectors as the control group, and get the data from 2019, one year before New Jersey rejoined RGGI, to 2021, one year after it rejoined.

(3) To investigate if there was any cross-border effect across states, the treatment states are either neighboring California or RGGI states. As GHGRP data have been available since 2010, we consider facilities in the states sharing the borders with California, namely, Arizona, Nevada, and Oregon, as the treatment group, and those that do not share borders with California and did not adopt any CAT programs as the control group. We take the data from 2010, three years before the treatment, to 2016, three years after the treatment.

Table 4: Difference-in-Difference Model for Research Question 2

	Pre treatment	Post treatment
Treatment: New Jersey power sector companies	2019 - 2020	2020- 2021
Controlled: New Jersey non-power sector companies	2019 - 2020	2020 - 2021

Table 5: Difference-in-Difference Model for Research Question 3

	Pre treatment	Post treatment
Treatment: companies in States adjacent to California	2010 - 2013	2013 - 2016
Controlled: Companies in states not adjacent to California and not in the RGGI	2010 - 2013	2013 - 2016

4. Outcome and Hypothesis

The outcome of each research question depends on whether the difference in GHG emissions between the treatment group and control group after the implementation of CAT programs changed significantly compared to the difference before the implementation. Our alternative hypotheses for the three questions are as follows, respectively.

(1) Ha: The multi-sector CAT program is better than single-sector programs at reducing GHG emissions or slowing growth In other words, in this case, even though economic sectors are dynamically connected, the crowd-out effect of single-sector programs, if any, is limited compared to multi-sector programs.

(2) Ha: Power-sector CAT programs affected the GHG-emission reduction or growth slowdown in non-power sectors. In this case, the superiority of multi-sector programs over power-sector

ones was possibly explained by the direct contribution of non-power sectors to GHG-emission reduction or growth slowdown.

(3) Ha: CAT programs had a cross-broader effect on the GHG-emission reduction or growth slowdown in surrounding states, which means the GHG emissions reduced or slowed in the increase in CAT adopter's neighbor states compared to non-neighbor states.

5. Threats

5.1 Concerns of clustering

We are concerned that, regardless of the research design, observations may cluster at the state and sector levels, weakening the power of our analysis.

For example, research question #1 studies the effectiveness of multi-sector CAT in California compared to single sector CAT in the states of RGGI. Companies in the same state may exhibit similar characteristics, such as attitudes toward low-carbon transition influenced by state political settings, financial conditions influenced by state financial policies, and the likelihood of developing or adopting clean energy technologies influenced by state R&D budgets. Too many similar companies tend to reduce sample variance and thus bias our estimation. Likewise, companies in the same industry may have similarities in business model, management style and financial status.

To mitigate the threat of clustering effect, we first calculated the minimum sample size required to ensure 80% power when observations cluster at the state and sector level versus when they do not with simulated data, and to ensure a valid analysis, we will survey samples of a larger size between the clustering and non-clustering cases.

The process of examining the minimum sample sizes for clustering and non-clustering cases is explained as follows. Based on the data availability and DiD design for research question #1, for both cases, we generated company-level data in 12 states and 9 sectors from 2010 to 2016, including a binary CAT intervention, a binary indicator of whether post-CAT intervention, company revenues, continuous energy efficiency, the amount of R&D investment, whether the company adopts new technologies, and the amount of GHG emissions.

For the clustering case, we incorporated random error terms at the state, sector, year, and company level. In the non-clustering case, we only incorporated them at the year and company level, with yearly variation being controlled by including year fixed-effects in the regression. We then simulated the data generation process for 200 times in a certain sample size until we found the sample sizes in which over 80% of coefficient estimates are significant at 5% statistical level. The results are listed in the table below:

Table 6. Minimum sample size with a power of 0.8

Total sample size	Sample size in each state, sector and year	Clustering	Power
1512	2	No	0.94
3780	5	No	0.94
7560	10	No	1
15120	20	Yes	0.79
18900	25	Yes	0.74
22680	30	Yes	0.8

According to Table 6, when there is no clustering in the data at the state and sector level, a sample size of 2 in each state, sector and year is sufficient to assure over 90% of power. When observations are clustered at the state and sector level, the sample size by state, sector and year has to exceed 30 to guarantee a power of 80%.

Therefore, to avoid the clustering effect, we will need at least 30 sample sizes in each stratum in each year to have adequate possibility to observe significance in our DiD model. Given the 30 sample size in each stratum of each year, we calculated the Minimum Detectable Effect (MDE) for the clustering case and non-clustering case. Results are presented below

Table 7: Minimum detectable difference-in-difference effects with sample size of 22680

Total sample size	Sample size in each state, sector and year	Clustering	DiD effect	Power
22680	30	No	1	0.59
22680	30	No	2	0.73
22680	30	No	3	0.86
22680	30	Yes	4	0.71
22680	30	Yes	5	0.79
22680	30	Yes	6	0.82

If companies are not clustered at the state and sector levels, the difference in GHG emissions between California and RGGI states would need to exceed 3 metric tons of CO2 to ensure a power of 80%.

For the clustering case, the MDE is 6, meaning that the difference in GHG emissions between California and RGGI states would need to exceed 6 metric tons of CO2 to ensure a power of 80%.

Suppose the DiD effect is 3 for non-clustering scenarios, the average GHG emissions simulated in the data generation process for California and RGGI states before the CAT intervention are 3939 and 3940 metric tons of CO2, respectively. For the clustering case, if the DiD effect is 6, the average GHG emissions generated for California and RGGI states before the CAT intervention are 5498 and 4023 metric tons of CO2, respectively. Therefore, a DiD effect of 3 or 6 is not significant compared to the average GHG emission magnitude, justifying our choice of sample size. (The means of stimulated emission data for California and RGGI states pre and post-treatment are reported in the Table A1)

5.2 Internal Validity

The internal validity of this study may be undermined by omitting unobserved variables that are correlating with CAT intervention and GHG emissions. Although we include a recall survey asking additional information to complement the GGRP data, some company attributes that are not observable or hard to measure by survey. Those left-behind CAT-related events may affect the control and treatment groups differently. If it's true, the difference between the treatment and control groups could be due to other factors instead of exposure to the CAT.

5.3 External Validity

This study also has potential problems in terms of external validity. The fact that only California has implemented multi-sector CAT limits the analysis to comparing California and RGGI states. Some distinct features of California companies might make the effect of multi-sector CAT unique to California. Therefore, it requires cautions to apply the implications of multi-sector CAT of this research to other cases.

Appendix

Table A1: The means of simulated GHG emission values for California and RGGI states pre and post-treatment (Metric Tons CO2)

	Pre treatment	Post treatment
Treatment: California companies	2010 - 2013 3939 (no clustering effect) 5498 (clustering)	2013 - 2016 3697 (no clustering effect) 5278 (clustering effect)
Controlled: companies in other 9 states	2010 - 2013 3940 (no clustering effect) 4023 (clustering)	2013 - 2016 3940 (no clustering effect) 4000 (clustering)

References

[1] Bartram, S. M., Hou, K., & Kim, S. (2022). Real effects of climate policy: Financial constraints and spillovers. Journal of Financial Economics, 143(2), 668–696. https://doi.org/10.1016/j.jfineco.2021.06.015

[2] Chan, N. W., & Morrow, J. W. (2019, May). *Unintended consequences of cap-and-trade? Evidence from the Regional Greenhouse Gas Initiative. Energy Economics, 80, 411–422. https://doi.org/10.1016/j.eneco.2019.01.007*

[3] Ehigiamusoe, K. U., & Lean, H. H. (2019). Effects of energy consumption, economic growth, and financial development on carbon emissions: evidence from heterogeneous income groups. *Environmental Science and Pollution Research, 26(22)*, 22611-22624. https://doi.org/10.1007/s11356-019-05309-5

[4] *Regional Greenhouse Gas Initiative (RGGI)*. (2019, December 5). Center for Climate and Energy Solutions. https://www.c2es.org/content/regional-greenhouse-gas-initiative-rggi/