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.

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

3.1 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. 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:

(1) How substantial the revenue generated by auctioning its allowances will be, especially for the countries/states, where their governments are in budget deficits.
(2) Empirical lessons in setting the 'price' of allowances in firms across sectors.
(3) 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 passed a law of multi-sector CAT and launched in 2013, covering virtually all economic sectors. Washington is the other that implemented a multi-sector CAT program, but it was just taking effect this year, 2023. As of today, eleven have launched the CAT programs under the Regional Greenhouse Gas Initiative (RGGI) since 2009, targeting the power sector only. New Jersey dropped out of RGGI in 2012 and rejoined in 2020.

Based on such background, the interventions differ for our three research questions:

(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. As the interventions started in the past, we entail retrospective observational GHG emission data to be able to evaluate the interventions regardless of possible methodologies. Given the availability of GHG emission data, we use facilities as the unit of analysis. We follow EPA to define a facility as all buildings, equipment, structures, and other stationery items located on a single site or adjacent or contiguous sites owned or operated by the same person (40 CFR Section 372.3). Therefore, the population of our study is all facilities in the U.S. across sectors. According to the places where interventions were implemented, our sample will be drawn from facilities in California, RGGI states, and neighboring states, including Arizona, Nevada, Oregon, Pennsylvania, West Virginia, Kentucky, and North Carolina. The sample size, time covered, and how to draw our sample will be determined by the availability of data, which will be explained in the following sections.

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. Facilities determine whether they must report based on the types of industrial operations located there, their emission levels, or other factors. Facilities are generally required to submit annual reports if:

(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.

Data Report and Verification

The Greenhouse Gas Reporting Program prescribes methodologies that must be used to determine GHG emissions from each source category. Reporters generally can choose among several methods to compute GHG emissions. The existing environmental monitoring systems and other factors may influence the decision about which method to use. All reports submitted to EPA are evaluated by electronic validation and verification checks. If potential errors are identified, EPA will notify the reporter.

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 facility data and emissions data.

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 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.

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 relative to 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.

(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.

$$Emission_{it} = \beta_0 + \beta_1 CAP_i + \beta_2 Post_t + \beta_3 CAP_i \times Post_t + \beta_4 Year_t + \varepsilon_{it}$$

Where CAP_i equals 1 if the facility is in the treatment group; $Post_t$ equals 1 if the year is after the implementation of CAP(2013); $Year_t$ is a continuous variable representing which year the emission was measured.

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 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. Threat

Data limitation:

(1) The GHGRP counts only large emitters. The conclusion drawn from this analysis of the data might have limited external validity.

Model concerns:

(2) Since California is the only state with a multi-sector program, the imbalance between the control and treatment groups in our research question 1 could result in imprecise estimation.
(3) The trends in emission between the treatment group and control group may be different prior to the CAT. We need to examine the trend in both groups before CAT was implemented using data from 2010-2013. If the trends are significantly different before the CAT, the Difference-in-Difference estimator could be biased.
(4) Unexpected events (unrelated to CAT) 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.