STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's Thesis in Economics
Academic Year 2020-2021

Can China's Carbon Emission Trading Make a Difference?

A quasi-experimental analysis of the ETS pilot scheme

Shicheng Xia (41653)

Abstract

With an aim to reduce carbon emissions and thus cope with the climate change, China initiated its Emission Trading System (ETS) in several pilot regions since 2013. In this thesis I investigate the CO_2 emission reduction effect of the China ETS pilot scheme, which could serve as an important indicator to infer how the forthcoming national ETS will perform in the future. Based on the city-level panel dataset of CO_2 emissions along with the city characteristics from 2003 to 2017, I apply the Difference-in-Differences (DID) estimator to capture the reduction of CO_2 emissions induced by the ETS pilot scheme in the participating cities. In order to tackle the latent confounding factors, a propensity score matching (PSM) strategy is also adopted. The results of PSM-DID model specification indicate that the ETS pilot scheme can significantly reduce the CO_2 emissions by around 8% in participating cities. Moreover, I find that there is no significant heterogeneity in ETS treatment effect across cities with different income levels. I further conduct placebo tests of changing ETS treatment timings and treatment group compositions as well as other sensitivity tests to check the robustness of the estimates. Additionally, I also investigate the impacts of the ETS pilot scheme on local economy and air quality in the pilot cities. No strong evidence is found to support the argument that the ETS would cause losses in jobs or outputs in the regulated industry.

Key words: China emission trading, carbon neutrality, policy analysis, quasi-natural experiment, matching

JEL: H23 Q51 Q54 Q58

Supervisor:	Pamela Campa
Date Submitted:	May 17, 2021
Date Examined:	May 25, 2021
Discussant:	Isak Thelander
Examiner:	Matilda Kilström

Acknowledgement

This thesis would not have been possible without the guidance and the help of many people.

Foremost, I would like to express my sincere gratitude to my supervisor, Pamela Campa, for her careful guidance and invaluable advice throughout the entire writing process. Besides, the ideas of my thesis were exactly inspired by her excellent lectures last year.

Special thanks goes to Dr. Matilda Kilström, Isak Thelander and Jeongeun Kim, for their insightful discussion and helpful comments about my thesis during the seminars. I am also thankful to all whom I have met and studied with at the Stockholm School of Economics, Sweden.

Ultimately, I am deeply grateful to my family, for their unconditional love and full support all the time.

Contents

List of Figures 3

List of Tables 4

1 Introduction 1

2 Literature Review 4

2.1 Emission Trading Policies . . . 4

2.2 The ETS Pilot Scheme and China's Environmental Issues . . . 5

3 A Profile of China ETS Pilot Scheme 6

3.1 Treated Regions and Treatment Time Period . . . 6

3.2 Map of the Distribution of Treated Regions . . . 8

4 Data Sources and Summary Statistics 9

4.1 Dataset and Variables . . . 9

4.2 Summary Statistics . . . 10

5 Quasi-Experimental Approach 14

5.1 Unconditional Difference-in-Differences . . . 14

5.2 Semiparametric Conditioning Strategy . . . 15

6 Empirical Results 17

6.1 DID Estimation Without Matching . . . 17

6.2 Propensity Score Matching . . . 19

6.3 PSM-DID Estimation Output . . . 22

6.4 Heterogeneity in Treatment Effect . . . 23

6.5 Robustness Checks . . . 25

6.5.1 Test for Parallel Trend Assumption . . . 25

6.5.2 Placebo Test Using Alternative Treatment Timing and Treated Units . . . 26

6.5.3 Check for the Uniqueness of the ETS Treatment Effect . . . 26

6.5.4 Cope with Serial Correlation in Treatment Variable: Data Aggregation Technique . . . 28

6.5.5 Test for the Stability of PSM-DID Estimates . . . 29

7 Other Impacts of the ETS Pilot Scheme 30

7.1 The Local Air Quality . . . 31

7.2 The Local Economy . . . 33

8 Concluding Remarks 36

References 38

Appendix 41

List of Figures

1 Participating Regions to the ETS Pilot Scheme in China 8
2 Trajectory Plot for Outcome Variables 12
3 Event Study: The Effect of ETS on City CO_2 Emissions 18
4 Nearest-Neighbor Matching Quality 20
5 Matching Quality: Kernel Distribution before/after NN Matching 21
6 Event Study: The Effect of ETS on Local CO_2 Emissions After Matching 25
7 Event Study: The Effect of ETS on Local PM2.5 Levels 32
8 Event Study: The Effect of ETS on Local Economy 35
9 Outcome of Stata Command pstest for NN Matching 41
10 Participating Provinces and Cities to the LCC Project 43
11 Outcome of Stata Command pstest for One-to-One Matching 46
12 Outcome of Stata Command pstest for Radius Matching 46
13 Outcome of Stata Command pstest for Kernel Matching 47
14 One-to-One Matching Quality 47
15 Radius Matching Quality 48
16 Kernel Matching Quality 48
17 Matching Quality: Kernel Distribution for Alternative Matching 49

List of Tables

1 Summary Statistics for Outcome Variables . . . 10
2 Mean Outcomes by Groups Before and After the Implementation of ETS . . . 11
3 Summary Statistics for Control Variables by ETS Treatment Group . . . 13
4 DIFFERENCE-IN-DIFFERENCES ESTIMATES WITHOUT MATCHING . . . 19
5 Nearest-Neighbor Matching Result . . . 19
6 Conditional Difference-in-Differences Regression Outcomes . . . 22
7 Results of Heterogeneity Analysis . . . 24
8 Placebo Tests For PSM-DID Model . . . 26
9 DID Estimation Including LCC Policy . . . 27
10 Regression Outcome Using Aggregated Data . . . 29
11 DID Regressions for Stability Test . . . 30
12 PM2.5 as the Outcome Variable . . . 31
13 Placebo Tests For PM2.5 Mitigation Effect . . . 33
14 Estimation for the Impact of ETS on Local Economy . . . 34
15 PSM-DID Regression Outcomes Using Time-Varying ETS Term . . . 42
16 DID Regressions Using Alternative Matching Algorithms . . . 45
17 One-to-One Matching Result . . . 45
18 Radius Matching Result . . . 45
19 Kernel Matching Result . . . 45

1 Introduction

As a landmark international accord that was adopted by nearly every nation, the Paris Agreement initiated since 2015 is committed to address climate change and its negative impacts. The agreement aims to substantially reduce global greenhouse gas emissions and thus limit the global temperature increase in this century. China, the biggest CO_2 emitting country on the earth, pledged in the Paris Agreement that it would propel an overall reduction of CO_2 emissions per unit of GDP by 40–45% below 2005 levels by 2020. The Chinese government has also set itself the goal of reaching the peak of CO_2 emissions by around 2030 and reducing CO_2 emissions per unit of GDP by 60 to 65% compared to 2005 levels [1]. In an effort to achieve these targets, China has embarked on establishing a national Emission Trading System (ETS) as one of the key building blocks of its carbon mitigation strategy. With the official work plan initially announced in 2011, China practically launched the pilot scheme for carbon trading market in several pioneer regions starting from the year of 2013. Once fully launched nationwide, the China ETS is projected to become the world's largest emission trading system covering one-seventh of global CO_2 emissions from fossil-fuel combustion [2]. Measuring the magnitude and depth, China's carbon neutrality policies play an important role in tackling the tremendous pressure in domestic environmental pollution as well as the global challenge of climate change.

Thereupon, this thesis studies mainly the impact of China's action on introducing the carbon emission trading system. As a specific type of market-based environmental policy instruments, emission trading policy, also known as cap-and-trade policy, has become increasingly popular as a tool for addressing a wide range of environmental issues among the major economies around the world. Market-based instruments provide incentives for the greatest reductions in pollution by those firms that can achieve these reductions most cheaply via markets, price, and other economic variables [3]. Theoretically, the properly designed and implemented market-based instruments can realize any desired level of emission abatement at the lowest overall cost to society (Stavins, 2003). While the traditional command-and-control (CAC) environmental policies require policy makers to gather more detailed information about the firms' compliance costs in order to achieve the same effectiveness, market-based instruments offer a more cost-effective approach for the government to control pollution. However, the performance and efficacy of marketed-based policies in real-world practice are still contentious. For example, some heavy polluters may choose to purchase emission allowances rather than to truly abate. And environmental injustice issues may arise with the potential unfair redistribution of regulated emissions caused by the market-based environmental policies. Specifically, under the context of China ETS pilot scheme, it deserves academic attentions in ex-post investigation of whether the emission trading policy really has the advantage of generating progressive environmental effects. Moreover, the assessment could also shed light on what the forthcoming nationwide ETS in China will deliver.

Following the announcement of the list of seven regional pilot ETS by the Chinese government, the rollout of the ETS pilot scheme commenced in 2013 and two provinces [4] as well as all four centrally-administered mu-

[1] https://chineseclimatepolicy.energypolicy.columbia.edu/en/unfccc
[2] https://www.iea.org/reports/chinas-emissions-trading-scheme
[3] https://www.epa.gov/environmental-economics/economic-incentives
[4] Namely the Guangdong province and the Hubei Province. Also notice that Shenzhen city pertained to Guangdong province was the first city to launch the ETS pilot scheme as early as June 2013.

nicipalities [5] in China had complied with the ETS pilot scheme by mid-2014. And in 2016, another province [6] launched its pilot ETS [7]. The assignment of the ETS treatment status was entirely determined by the central government and no official document about selection criteria was published, therefore, the ETS pilot scheme cannot be viewed as a strictly randomized experiment but rather a quasi-experiment [8]. Within each complying region, the government sets the cap/limit of total CO_2 emissions [9] for facilities in the regulated sectors, and create a market for the emission allowances which can be traded among polluting firms. The purpose of the pilot ETS practice was to encourage the complying cities and provinces to explore the optimal design for the carbon trading mechanisms and promote their experiences for the establishment and implementation of the future national ETS which the pilots themselves will eventually be merged into. Therefore, it is expected that the market-based instrument of carbon emission trading could generate desirable CO_2 abatement effect in these participating regions before expanding nationwide. Nevertheless, skepticism remains about whether the ETS pilot scheme is truly effective despite the fact that the authority in China claimed the success of it after years of implementation and enacted the work plan for the national ETS in December 2017. The most essential question is: can the ETS pilot scheme effectively reduce CO_2 emissions beyond what could have been achieved with more prescriptive CAC regulation? Meanwhile, other concerns about the ETS pilot scheme are also noteworthy: would it be accompanied with high cost to the local economy? Would it induce unfair redistribution of pollution across different regions?

To examine these issues, the empirical approach in the thesis is based on a quasi-experiment that exploits two sources of variation in the introduction of the China ETS pilot scheme. Amongst all 31 first-level administrative regions [10] in China, the ETS pilot scheme has been operating in 7 of them since 2013. To attain the primary goal of identifying the causal effects of carbon emission trading, I mainly adopt the Difference-in-Differences (DID) estimator that compares the outcomes of interest, especially the CO_2 emissions, in the ETS participating cities and non-participating cities before versus after 2013. The data range for the empirical analysis is from 2003 to 2017. Specifically, due to the intrinsic non-randomized nature of the ETS treatment assignment, the credibility of causal inferences is fostered by the construction of tenable counterfactual estimate of the outcomes. A propensity score matching (PSM) strategy is thus employed along with the DID estimator in order to mitigate the potential confounding factors and extract the policy impacts.

The results suggest that on average an approximately 8 percent decline in CO_2 emissions for the complying cities vis-à-vis the control group was yielded by the China ETS pilot scheme. Placebo tests are conducted to check the robustness of the results and no significant policy effect is found before the actual outset of the ETS pilot

[5] Formally the municipalities under the direct administration of central government. They are of the highest level of classification for cities used by the People's Republic of China. Municipalities have the same rank as provinces, and form part of the first tier of administrative divisions of China. A municipality is a "city" with "provincial" power under a unified jurisdiction. As such it is simultaneously a city and a province of its own right (Wikipedia).

[6] Namely the Fujian province.

[7] https://ets-china.org/ets-in-china/

[8] Nevertheless, as stated by Greenstone and Gayer (2009), valid causal inference that the difference in the outcome variables is due to the explanatory policy can still be drawn by controlling some important covariates, balancing the pre-treatment observable characteristics across the treatment and control groups, or applying the instrumental variable (IV) approach.

[9] Except for the ETS pilot scheme in Chongqing, which also regulates several other GHG emissions such as CH4 and N_2O. Other pilot regions target only at CO_2 emissions.

[10] Excluding the two Special Administrative Regions, Hong Kong and Macao, and the disputed Taiwan Province under separate rule (Wikipedia).

scheme nor in the factitious treatment group constituted by non-complying cities. In further investigation about the unconfoundedness assumption, I show that estimation of the pilot scheme's CO_2 emission reduction effect remains salient when the CAC-type environmental policy targeted at the carbon emissions, that is, the Low-Carbon City (LCC) project, is accounted for. Moreover, two hypotheses related to the stability of the estimating results, or the Stable Unit Treatment Value Assumption (SUTVA), namely: 1) CO_2 emissions in the ETS participating cities tend to flow to the nearest non-participating cities; 2) CO_2 emissions in the ETS participating cities tend to flow to cities with lax environmental regulations, are tested and I reject both of them. Furthermore, there is no evidence for the existence of heterogeneous ETS treatment effect across participating cities of different income levels.

In addition to identifying the CO_2 emission reduction effect, I also evaluate the potential impacts of the ETS pilot scheme on the local air pollution and economy. These two aspects are represented by surface PM2.5 concentration levels and performance of regulated industries respectively. The results of estimating the PM2.5 mitigation effect are ambiguous due to the interference of other relevant policies or events. Meanwhile, no strong evidence can be revealed to support the statement that the ETS pilot scheme would jeopardize the average employment or outputs of the regulated industries in treatment group, although the cities with higher reliance on the secondary sector seem to suffer more losses in industrial output than others. For a number of reasons, especially the limitations of the city-level aggregate data used in the thesis, the results of this part may not perfectly capture the nuanced performance of the firms in the regulated industries and therefore need to be interpreted with caution.

Via the empirical analysis for the ETS pilot scheme in China, this thesis illuminates its efficacy of reducing carbon emissions and thus highlights that China's adoption of market-based environmental policies can generate progressive environmental impacts and ultimately contribute to the global effort to combat climate change. Besides, several improvements upon the existing literature about the China ETS are made by this thesis. Firstly I make use of the city-level panel data that can exploit more sufficient variations compared to the prior related studies usually conducted at province-level (Zhang et al., 2019a; Dong et al., 2019; Zhang et al., 2020; Qi et al., 2021). Also the satellite imagery CO_2 emission data from China Emission Accounts & Datasets (CEADs) [11] used in the thesis provide a more accurate and extensive measurement of China's regional CO_2 emissions than the common practice adopted by the existing literature of reckoning the CO_2 emitting amount from the industrial output data. Moreover, I discuss the potential heterogeneity issues as well as the influence on local socioeconomic status from the ETS pilot scheme, which are rarely considered in the previous researches. Last but not least, a series of robustness checks are designed to enhance the validity of identifying the causal relationships between the ETS pilot scheme and the outcomes of interest in the thesis.

The remainder of the thesis is organized as follows: Section 2 reviews the findings from the existing relevant literature. In Section 3 and Section 4 I provide a brief profile about the China ETS pilot scheme and present the descriptive summary for the data used in the empirical analysis, respectively. Section 5 introduces the methodology and empirical approach. Subsequently, in Section 6 and Section 7 I demonstrate the empirical results and give interpretations. Finally in Section 8 I draw the conclusions and discuss.

[11] https://www.ceads.net/

2 Literature Review

2.1 Emission Trading Policies

Before the trend of transiting towards the more market-oriented instruments for environmental regulation, policy makers mainly adopted the command-and-control type, or CAC policy to regulate the emissions in the economy. By definition, CAC policy relies on regulation (permission, prohibition, standard setting and enforcement), while market based instruments are centered around providing financial incentives to internalise the environmental externalities (OECD, 2008). As argued by Stavins (2003), due to the varying costs of controlling emissions among firms, and even among sources within the same firm, the CAC-type policies lose flexibility and efficacy as they often set uniform standards. Futhermore, CAC regulations tend to freeze the development of technologies that might otherwise result in greater levels of control, as little or no financial incentive exists for businesses to adopt new technology and exceed their control targets. In contrast, market-based instruments can provide powerful incentives for companies to adopt cheaper and better pollution-control technologies. Specifically, in the emission trading programs, the governments set the cap/limit for total emission, and create a market for emission permits which can be traded among polluting firms. Within the cap, firms receive or purchase emission permits so that in equilibrium all firms operate at the optimal level where firm's marginal abatement cost equals the market price for emission. Typical examples for the practice of emission trading include the 1990 Clean Air Act Amendments (CAAAs) of the United States which initiated the transition from CAC policies towards emissions trading policies [12], and the 2005 Emissions Trading System of European Union (EU ETS) which imposes an overall cap on CO_2 emissions in 31 countries and is now the largest cap-and-trade system worldwide [13].

The empirical analyses about the performance of the market based policy, especially the widely adopted emission trading policy, have been conducted luxuriantly in the economic literature. A considerable quantity of evidence is provided to testify the efficacy of emission trading programs around the world. Under the context of the largest emission trading system so far, the EU ETS, Bayer and Aklin (2020) base on a generalized synthetic control method and estimate the treatment effect of the program. They argue that the EU ETS brings a decrease of CO_2 emissions in regulated sectors of between 8.1% and 11.5% against the counterfactual. Fowlie (2010) estimates the NO_X emission reduction effect of the NO_X Budget Program (NBP) in the eastern and midwestern United States between 2003 and 2008. She finds that the program delivers a 10% to 14% lower NO_X emission for the regulated facilities, though the heterogeneity in electricity market regulations causes a larger share of the permitted pollution to be emitted in states with more severe air quality problems. On the same object, Deschênes et al. (2017) investigate the impacts of the NBP on emissions as well as the residents' medical spending to quantify the benefits to health. They show that the program performs well and makes positive contribution on both aspects.

Albeit the desirable traits of market-based environmental policies like ETS, a major concern for them lies on the environmental injustice issues. As heavily polluting facilities can purchase large quantities of emission permits rather than truly abate, the permitted emissions may flow disproportionately to the disadvantaged areas.

[12] https://www.epa.gov/clean-air-act-overview/1990-clean-air-act-amendment-summary
[13] https://ec.europa.eu/clima/policies/ets/

An empirical example is offered by Hernandez-Cortes and Meng (2020). They argue that the California's carbon market induced changes in the spatial distribution of local air pollution. They further integrate econometric model with pollution transport model to estimate the environmental justice (EJ) gap and conclude that carbon market widened the EJ gap in NO_X, SO_X, PM2.5, and PM10 prior to 2013 and since then decreased the EJ gap by 21-30% across pollutants. Ringquist (2011) finds that the Clean Air Act Amendment's (CAAA) sulfur dioxide allowance trading program (ATP) in the US appears to concentrate SO_2 emissions in communities having large percentages of adults without a high school diploma. On the contrary, Corburn (2001) suggests that for the first few years of implementation the Acid Rain Programme (ABP) in US does not appear to have been concentrating SO_2 pollution disproportionately for the poor and racial minority populations. Similarly, Fowlie et al. (2012) investigate the potential environmental justice issues in the Southern California's cap-and-trade program, RECLAIM. They show that average NO_X emissions fell 20% at RECLAIM facilities relative to the counterfactual and this effect does not vary significantly with neighborhood demographic characteristics.

2.2 The ETS Pilot Scheme and China's Environmental Issues

As the China ETS pilot scheme has been implemented for several years and a national level carbon trading system is around the corner, recently the studies about the pilot scheme's carbon emission reducing effect, as well as its impacts on other aspects such as industrial production or green technology have been emerging with higher frequency.

The existing literature mostly share a positive attitude towards the China ETS pilot scheme when identifying its emission reduction effect. Various identifying approaches are adopted in relevant empirical analyses. Liu et al. (2017) use a Chinese multi-regional general equilibrium model (TermCO2) to study the ETS effect in one of the participating provinces and show that in 2014 the ETS pilot brings down the carbon emissions by 1.00% while causing a 0.06% GDP loss. Zhang et al. (2020) test the impact of ETS on the carbon emission reduction and economic growth using DID estimators and conclude that the implementation of ETS can reduce the carbon emissions by 24.2% in the pilot provinces and increase the gross industrial output value by 13.6%. Zhang et al. (2019a) find that the ETS pilot has a smaller emission reducing effect of 10.1% and can decrease the industrial carbon intensity by 0.78%. Zhou et al. (2019) study the effect of ETS pilot on carbon intensity of the pilot provinces, suggesting an annually decline of approximately 0.026 tons per 10000 CNY of GDP. Tang et al. (2021) analyze the provincial industrial-level data and argue that the pilot ETS can effectively reduce the regulated industries' carbon emissions by about 8% on average while this effect has a substantial heterogeneity for different pilot provinces and industries.

In contrast to the papers that put forward positive findings about the effectiveness of China ETS pilot scheme, the literature sharing opposite views usually argue about the unsustainable CO_2 reducing effect of the policy, heterogeneity in the policy's impact across different regions, or the potential costs. Dong et al. (2019) compare the effects in economic output and carbon emissions between ETS pilot and the command-and-control type of policies and find no significant difference. Additionally, they highlight that the CAC policies can achieve the same emission reduction with less economic loss comparing to an inefficiently-operated ETS. Shen et al. (2020) study the effect of China's ETS at the firm-level by focusing on the policy executors within the regulated industries and find the

CO_2 emission reduction effect of the policy to be heterogeneous across firms of different sizes, and moreover, the effect tends to attenuate over time.

Notably, before the introduction of the ETS pilot scheme, there are also other environmental policies targeted at CO_2 emissions. However, most existing literature discussing the China ETS did not consider them. An example for the prior carbon reduction efforts made by China is the command-and-control policy of Low-Carbon City program (LCC). Regarding central government's guidelines about the low-carbon urbanization, local governments of the LCC complying cities conduct city-wide low carbon planning and set up city-level energy and emission targets. They also receive fiscal and policy support from the central government. Fu et al. (2021) do a retrospective study to evaluate the impacts of LCC policy on carbon emission efficiency of Chinese cities and argue that it significantly improves the carbon emission efficiency of the treatment group while it might take longer time for the policy to improve the carbon emission efficiency than to reduce the absolute amount of carbon emissions.

The CO_2 emissions from human's manufacture and life also interplay with the local criteria air pollutants. Specifically, in China the fine particulate matters (PM2.5) have been significantly affecting the health and the quality of life for residents. Some studies regarding the relevant environmental issues are done in the existing literature. Xu et al. (2019) use mobile phone and census data to infer residents' wealth level and daily PM2.5 exposure in Beijing. They discover that in winter the commuters with low wealth level are exposed to 13% more PM2.5 per hour than those with higher wealth level. And Zhang et al. (2019b) discover that although economic development significantly reduces the health risks of PM2.5 in China, the reduction effect in the less-developed western regions is significantly less than that in the eastern and central regions, as well as at the national level, due to the smaller investment amount in science and technology. However, to the best of my knowledge, the relations between the China ETS pilot scheme and the local PM2.5 levels are rarely discussed. The only study investigating this issue I could find so far is conducted by Cao et al. (2021). Focusing on the synergistic effect of the ETS on local air quality, they simulate the potential effect of ETS pilot on air pollution in Hubei province using GIS-based spatial modeling, and find that although the overall level of PM2.5 concentration goes down, an increase is observed after the implementation of ETS in some areas of major cities in the province.

3 A Profile of China ETS Pilot Scheme

3.1 Treated Regions and Treatment Time Period

For each regional ETS pilot, I will briefly introduce its basic situations such as the launch date and the coverage of greenhouse gas (GHG) emissions. In general, the assignment of ETS treatment status is made at provincial level [14] . These ETS pilot regions also spread across all geographical divisions in the land of China (With Chongqing in Western China, Hubei province in Central China and the rest in Eastern China). All information for the ETS pilots comes from the ETS pilot factsheets published on the official website of the Sino-German project "Capacity

[14] Except for the ETS pilot in Shenzhen city. Note that the 4 municipalities, namely Beijing, Tianjin, Shanghai and Chongqing, are of the same administrative rank, that is, the first tier in the hierarchy of administrative divisions in China, as the other provinces.

Building for the Establishment of Emissions Trading Schemes (ETS) in China" [15].

Shenzhen ETS [16].—Launched on 18th June 2013, the regulated GHG is CO_2 and more than 40% of gross emissions are covered. 824 enterprises (in 2016) in sectors of power, water, gas, manufacturing, buildings, and transportation (port, subway, buses) are covered.

Shanghai ETS.—Launched on 26th November 2013, the regulated GHG is CO_2 and more than 57% of gross emissions are covered. 298 enterprises (in 2017) in sectors of airports, aviation, chemical fiber, chemicals, power and heat, water suppliers, hotels, iron and steel, petrochemicals, ports, shipping, non-ferrous metals, building materials, paper, railways, rubber, textiles and some service sectors are covered.

Beijing ETS.—Launched on 28th November 2013, the regulated GHG is CO_2 and more than 45% of gross emissions are covered. 943 enterprises (in 2017) in sectors of electricity, heating, cement, petrochemicals, other industrial enterprises, manufacturers, service sector and public transport are covered.

Guangdong ETS.—Launched on 19th December 2013, the regulated GHG is CO_2 and more than 60% of gross emissions are covered. 296 enterprises (in 2017) in sectors of power, iron and steel, cement, papermaking, aviation and petrochemicals are covered.

Tianjin ETS.—Launched on 26th December 2013, the regulated GHG is CO_2 and 55% of gross emissions are covered. 109 enterprises (in 2017) in sectors of heat and electricity production, iron and steel, petrochemicals, chemicals, oil and gas exploration are covered.

Hubei ETS.—Launched on 2nd April 2014, the regulated GHG is CO_2 and more than 35% of gross emissions are covered. 344 enterprises (in 2017) in sectors of power and heat supply, iron and steel, nonferrous metals, petrochemicals, chemicals, chemical fiber, cement, glass and other building materials, pulp and paper, ceramics, automobile and general equipment manufacturing, food, beverage and medicine producers are covered.

Chongqing ETS.—Launched on 19th June 2014, the regulated GHG is CO_2, CH4, N_2O, HFCs, PFCs and SF_6. More than 40% of gross emissions are covered. 237 enterprises (in 2016) in sectors of power, electrolytic aluminum, ferroalloys, calcium carbide, cement, caustic soda, iron and steel are covered.

Fujian ETS.—Launched on 30th September 2016, the regulated GHG is CO_2 and more than 60% of gross emissions are covered. 277 enterprises (in 2016) in sectors of electricity, petrochemical, chemical, building materials, iron and steel, nonferrous metals, paper, aviation, and ceramics are covered.

[15] https://ets-china.org/downloads/

[16] The city is pertained to Guangdong province.

3.2 Map of the Distribution of Treated Regions

FIGURE 1: PARTICIPATING REGIONS TO THE ETS PILOT SCHEME IN CHINA

Notes : The green zones are the ETS pilot provinces and the red spots are the pilots in municipalities and Shenzhen city. The map is from the official website of the Sino-German project "Capacity Building for the Establishment of Emissions Trading Schemes (ETS) in China". For details see https://ets-china.org/ets-in-china/.

4 Data Sources and Summary Statistics

4.1 Dataset and Variables

To facilitate the empirical analysis in the thesis, I collect data for the variables of interest from several different sources of data files and compile them into one set of balanced panel data. The dataset covers the observations for nearly all city-level administrative divisions in China [17] with the time period ranging from 2003 to 2017. In this section I describe the specific data sources and formats for both outcome variables (annually CO_2 emissions, surface PM2.5 concentration levels, power sector employment and industrial outputs) and control variables (divided into city characteristics and weather conditions).

CO_2 Emissions.—The primary outcome variable of interest is the CO_2 emissions in China. However, the Chinese authorities do not provide any official statistics for local level CO_2 emissions. Relevant carbon emission levels in the extant body of literature are usually estimated from local industrial outputs or energy usage. In an attempt to overcome the measurement errors and seek for more accurate measurement for local CO_2 emission levels, I use the latest dataset of the county-level CO_2 emissions [18] provided by China Emission Accounts & Datasets (CEADs) [19]. Note that county is the most basic governmental unit in China [20] so this dataset could better capture the regional variations. I further integrate the CO_2 emission data to city-level for compatibility with other variables. A city's annually CO_2 emission amount (measured by million tons) is calculated by summing up the emission quantities of all its subordinate counties.

Ground PM2.5 Concentration Levels.—As the Chinese government had not listed PM2.5 as an air quality criteria pollutant nor had it issued the relevant data to the public until 2013, to fulfill the research need in the thesis I use the China historical PM2.5 concentration data estimated by the Atmospheric Composition Analysis Group from the Washington University in St. Louis [21]. The dataset contains PM2.5 concentration levels (measured by Air Quality Index, or AQI [22]) from 2000 to 2018 and is converted to match each city and year.

Power Sector Employment and Industrial Outputs.—To measure these two outcomes for each city and year, I collect the official data for them from the city statistics yearbooks published by the National Bureau of Statistics

[17] Out of a total of 297 cities (without Hong Kong and Macao) in China, 284 are contained in the dataset, including the 4 municipalities. The uncovered cities are mostly from border provinces like Xinjiang and Tibet, or are newly-established cities like the Sansha city on islands in the South China Sea. Open official statistics in China reveal little information for these cities. Nevertheless, the "mystique" of these cities also reflects that they are remote from the economic centers and sparse in population. Thus I argue that it is reasonable to believe that the missing observations of them are not likely to affect the empirical results.

[18] Specifically, a particle swarm optimization-back propagation (PSO-BP) algorithm was employed to unify the scale of DMSP/OLS and NPP/VIIRS satellite imagery and estimate the CO_2 emissions in 2,735 Chinese counties during 1997–2017.

[19] https://www.ceads.net/data/county/

[20] Formally the county-level division is the third level of the Chinese administrative hierarchy, lying beneath city-level and province-level. Apart from normal counties, it also contains autonomous counties, county-level cities, banners, autonomous banners and City districts. (Wikipedia)

[21] The Atmospheric Composition Analysis Group produces combined geophysical-statistical estimates of PM2.5 over China using the recently expanded PM2.5 measurement network in this region from May 2014 to December 2018, and extends these values back to 2000 using the interannual changes between the GM observed and non-GM observed time periods based on satellite-derived values of Hammer et al. (2020). The dataset is published on their official website:https://sites.wustl.edu/acag/datasets/surface-pm2-5/

[22] The original PM2.5 concentration levels are measured by $\mu g/m^3$ and then coverted to AQI levels using the EPA standard. For details see: https://www.epa.gov/data-standards

of China (NBSC) [23]. Specifically, the power sector employment indicator counts the number of employed persons in the production and supply of electricity, heat, gas and water and belongs to the chapter of "Employment and Wages", while the industrial outputs indicator comes from the chapter of "Industry".

City Characteristics Vector.—I gather data for a series of relevant local socioeconomic indicators to reflect the city characteristics from the previously mentioned city statistics yearbooks provided by the National Bureau of Statistics of China. These indicators in the vector might be used as control variables or sources of heterogeneity in the empirical study. The following variables are contained: population, GDP, GDP per capita, total electricity consumption, amount of foreign direct investment (FDI), amount of investment in research and development (R&D), ratio of secondary industry in GDP, ratio of tertiary industry in GDP, afforestation rate and average wage for workers.

Weather Conditions Vector.—Other than city characteristics, variations in local weather conditions may also interact with the CO_2 emissions and PM2.5 concentrations. Therefore meteorological data are collected from the China Meteorological Data Service Center (CMDC) [24] to show the weather conditions in the cities. I compile the annual average levels of temperature, humidity, precipitation and sunshine duration for each city.

4.2 Summary Statistics

TABLE 1: SUMMARY STATISTICS FOR OUTCOME VARIABLES

	Observations	Mean	SD	25th Pct	Median	75th Pct
CO_2 Emissions (million tons)	4240	24.945	23.085	10.182	18.125	32.602
PM2.5 Levels (AQI)	4240	43.834	19.097	29.503	41.256	57.591
Industrial Outputs (10000 CNY)	4240	2.39×10^7	3.84×10^7	3.82×10^6	1.08×10^7	2.56×10^7
Power Sector Employment (persons)	4240	10531.9	9813.003	5000	8100	12500

Notes : In cells I report the number of observations, means, standard deviations, 25th percentiles, medians and 75th percentiles for the outcome variables of interest. Data are for the entire sample of cities and the whole time period from 2003 to 2017.

Table 1 summarizes the basic statistics for the outcome variables in my analysis. To further show the change in the outcomes induced by the ETS pilot scheme, I partition the data into four categories by treatment status of the cities and ETS treatment time. The year of 2013 is used to divide the time period before/after the introduction of ETS pilot scheme. Table 2 shows the group mean outcomes before and after the ETS implementation and the corresponding rate of change. In both groups there is an increase in the average CO_2 emissions, nevertheless, we can straightforwardly observe that the average CO_2 emission expansion rate in the treatment group is much lower than that in the control group, providing an initial hint that the ETS pilot scheme could be effective. Similarly, the

[23] http://www.stats.gov.cn/english/Statisticaldata/AnnualData/
[24] http://data.cma.cn/en

TABLE 2: MEAN OUTCOMES BY GROUPS BEFORE AND AFTER THE IMPLEMENTATION OF ETS

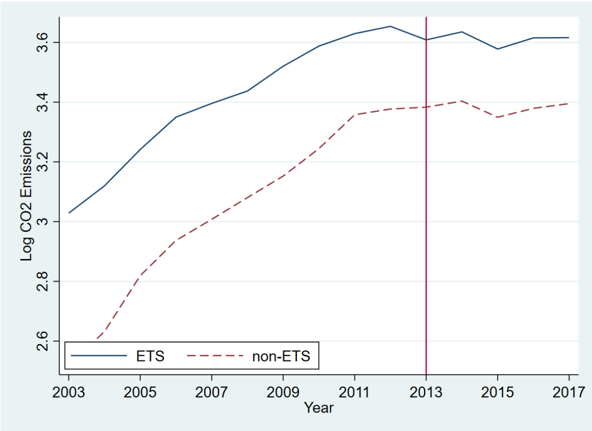
Group	Treatment Group			Control Group		
Time Period	Before	After	Rate of Change	Before	After	Rate of Change
CO_2 Emissions (million tons)	27.884	34.728	24.545%	21.101	29.438	39.509%
PM2.5 Levels (AQI)	42.996	36.093	-16.055%	45.506	42.241	-7.175%
Industrial Output (10000 CNY)	3.01×10^7	6.61×10^7	119.601%	1.39×10^7	3.33×10^7	115.827%
Power Sector Employment (persons)	13658.05	14026.42	2.697%	9446.183	10698.06	13.253%

Notes : In cells I compare the means for the outcome variables of interest in treatment/control group and before/after the start of ETS pilot scheme and calculate the rate of change. For convenience I ignore the different launch timings among ETS-participating regions and use the year of 2013 as the uniform cut-off in the time period to define the outset of ETS pilot scheme.

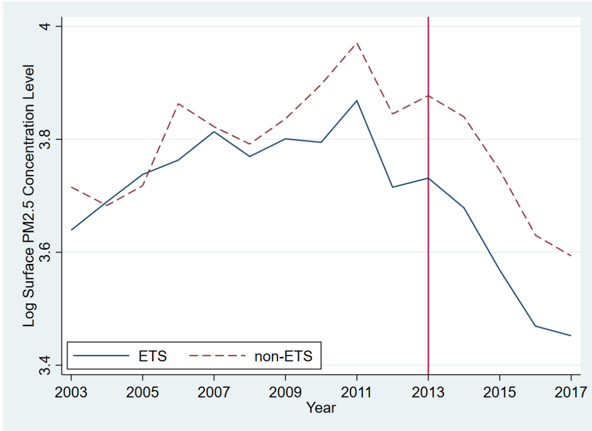
ETS treated cities also on average experience a greater decline in the PM2.5 concentration levels, which inspires the investigation of the possible PM2.5 mitigation effect of ETS pilot scheme. The growth rates in industrial outputs before and after 2013 are close between the two groups, although the average output in ETS treated cities is around twice as much as that of control group cities. The treated cities also have more employees in the regulated power sector and the size is quite stable, while in the control group the number of power sector employees increases considerably after 2013.

To provide a more dynamic view for the potential effect of ETS pilot scheme on the outcome variables of interest, in Figure 2 I demonstrate the trajectories of the outcome variables by group over the whole time period studied in this thesis. By inspecting Figure 2a, we can find some preliminary evidence for the existence of parallel trend between the CO_2 emissions of treatment group and control group before the year of 2013. And exactly in the year of 2013 the average CO_2 emissions in the treatment group began to drop. This motivates the adoption of the Difference-in-Differences approach, for which the underlying assumption is that the treated units and the controlled should have parallel paths of the outcome before the concerned policy comes into force, to study the treatment effect of the ETS pilot scheme. Of course this is just a rough visual inspection for the feasibility of this method and more rigorous quantitative analyses are required to examine the underlying assumptions. I will turn to this in the latter sections of the thesis. Things are less clear when inspecting the trajectory plots for the other outcomes. In Figure 2b, I find that on average both treated and untreated cities peaked their surface PM2.5 concentration levels in 2011 and experienced a drastic decrease immediately afterwards. This might be induced by policies or events other than the ETS pilot scheme and I will discuss about this in more detail later. Overall the downward trend maintained in the following years and no significant change seems to have occurred in 2013. Similarly, in the trajectories for industrial output and power sector employment, the outset of ETS pilot scheme

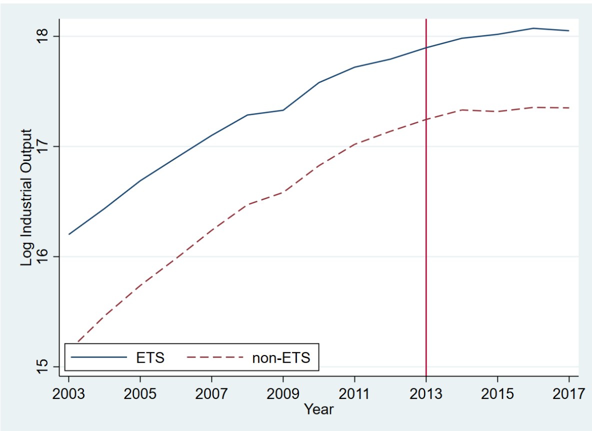
FIGURE 2: TRAJECTORY PLOT FOR OUTCOME VARIABLES



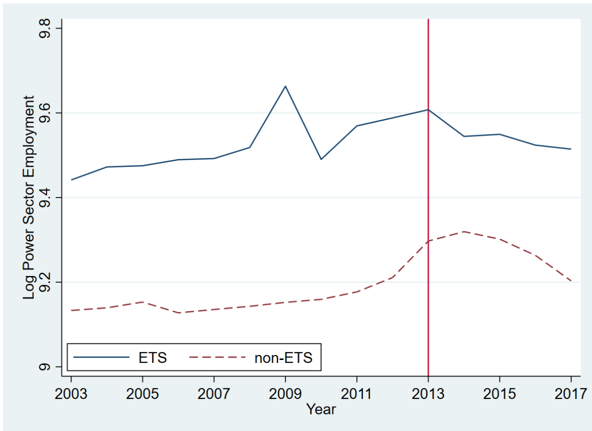
(A) CO_2 EMISSIONS TRAJECTORY



(B) PM2.5 CONCENTRATION LEVELS TRAJECTORY



(C) INDUSTRIAL OUTPUT TRAJECTORY



(D) POWER SECTOR EMPLOYMENT TRAJECTORY

does not appear to cause much deviations between the two groups either. Nevertheless, I will draw the conclusions via more precise analyses for this issue in the latter sections.

Additionally, in Table 3 I show the summary statistics for other relevant variables used in the empirical analyses of this paper. They are classified into 2 vectors, namely the city characteristics and the city weather conditions. The statistics are summarized for ETS treatment group and control group respectively.

TABLE 3: SUMMARY STATISTICS FOR CONTROL VARIABLES BY ETS TREATMENT GROUP

	Observations	Mean	SD	25th Pct	Median	75th Pct
Panel A. City Characteristics by Treatment Status						
Population	*Treatment* : 591	529.0524	629.0601	284.12	355	623.95
(10000 people)	*Control* : 2927	434.7006	338.3954	237.9	379.52	572.82
GDP	*Treatment* : 591	2.41×10^7	3.84×10^7	5404494	1.01×10^7	2.14×10^7
(10000 CNY)	*Control* : 2927	1.34×10^7	1.61×10^7	3967576	8133300	1.61×10^7
GDP per capita	*Treatment* : 591	5.149052	7.095185	1.314028	2.570649	5.508965
(10000 CNY)	*Control* : 2927	3.258597	3.218589	1.200774	2.295033	4.130367
Secondary Industry	*Treatment* : 591	49.17892	29.13873	41.65999	47.21	53.96
(%)	*Control* : 2927	48.79314	11.56277	41.82001	49.3	55.76
Tertiary Industry	*Treatment* : 591	41.10798	23.22267	33.87001	37.21	42.73
(%)	*Control* : 2927	37.11962	10.28635	31.17	35.85	41.56001
FDI Amount	*Treatment* : 591	129772.3	288485.3	10106	21766	118030
(10000 CNY)	*Control* : 2927	56840.38	139157.5	3686	13556	48767
R&D Investment	*Treatment* : 591	556741.8	1059263	99692	240367	511799
(10000 CNY)	*Control* : 2927	320436.3	339503.5	90247	213044	435735
Electricity Consumption	*Treatment* : 591	1494467	2388309	207874	441351	1419567
(10000 KW/h)	*Control* : 2927	653343.3	905822.7	154216	348170	780792
Afforestation Rate	*Treatment* : 591	41.35712	29.51085	35.03	39.86	42.72
(%)	*Control* : 2927	36.3857	9.215479	32.26	38.41	41.78
Wage	*Treatment* : 591	35568.53	21049.94	18460.14	30769.45	48294.13
(CNY)	*Control* : 2927	33902.2	18723.26	18317.41	31025.3	46999
Panel B. City Weather Conditions by Treatment Status						
Temperature	*Treatment* : 591	20.19509	2.781944	17.6	21	22.6
($°C$)	*Control* : 2927	14.37874	4.8132	11.2	15.4	17.7
Humidity	*Treatment* : 591	74.03655	5.459244	72	75	77
(%)	*Control* : 2927	67.59644	9.112835	62	69	75
Precipitation	*Treatment* : 591	1495.644	488.6023	1158	1441.3	1837.6
(mm)	*Control* : 2927	1958.301	31971.84	563.1	857.8	1235
Sunshine	*Treatment* : 591	1751.599	281.3145	1583.4	1756.7	1929.6
(h)	*Control* : 2927	1975.605	515.3467	1627.4	1970.2	2358.1

Notes : In cells I report the number of observations, means, standard deviations, 25th percentiles, medians and 75th percentiles for indicators in both city characteristics vector and city weather conditions vector. Data for all variables cover the whole time period from 2003 to 2017 and are classified by treatment group and control group. City characteristics data are from NBSC, while the weather conditions data are from CMDC.

5 Quasi-Experimental Approach

The feature of the ETS pilot scheme is that only cities of several province-level administrative divisions in China are selected to be exposed to the treatment of the policy, while the rest receives no treatment. This can be viewed as a quasi-experiment in which the assignment of treatment status is determined by factors beyond the researcher's control such as socioeconomic conditions, politics or nature. Therefore, it is suitable to apply the quasi-experimental techniques on the analyses of the ETS pilot scheme. As highlighted by Greenstone and Gayer (2009), even though the assignment is not random, valid inferences can still be drawn from the comparison between the outcomes of treatment and control groups if the assumption that group assignment does not relate to other determinants of the outcomes is satisfied. Especially, the essential target of this thesis is to estimate the environmental effects of China's ETS pilot scheme, and the empirical models I employ to do the causal inference are based on the DID approach. The main idea is to exploit the variations in the outcome variables of interest between the ETS pilot cities and the non-ETS cities as well as between the years before the implementation of ETS pilot scheme and the years after it. In this section, I will discuss what my identification strategies and underlying assumptions are, along with the corresponding econometric model specifications.

5.1 Unconditional Difference-in-Differences

Based on the ETS treatment status of each city, I let the treatment indicator $D_i = 1$ for city i if it is complying with the ETS pilot scheme, and let $D_i = 0$ if not regulated by the policy during the time period of the study. The outcomes of interest $Y_{it}(1)$ and $Y_{it}(0)$ represent the CO_2 emissions of city i in year t conditional on participation and nonparticipation to the ETS, respectively. Thus the average treatment effect on the treated (ATT) can be denoted as:

$$\alpha_{TT} = E[Y_{it'}(1) - Y_{it'}(0)|D_i = 1] \tag{1}$$

where t' represents the years after the implementation, and α_{TT} measures the average CO_2 emissions reduction effect of the policy on the ETS pilot cities.

Since the China ETS pilot scheme resembles a quasi-natural experiment with the policy treatment assignment being non-random, there are pitfalls detrimental to the validity of the ATT estimator. To detect the flaws in the ATT estimation, we can first look at the specification of the average treatment effect (ATE) shown in the equation below:

$$\alpha_{TE} = E[Y_{it'}(1) - Y_{it'}(0)|D_i = 1] + \{E[Y_{it'}(0)|D_i = 1] - E[Y_{it'}(0)|D_i = 0]\} \tag{2}$$

where in addition to the first term representing the ATT that I want to estimate, the second term nonetheless contains the selection bias that undermines the consistency of the ATT estimation. As noted by Greenstone and Gayer (2009), it measures the difference in potential untreated outcomes between the units that did and did not receive the treatment. Thereby, in order for the specification in Equation (1) to unbiasedly estimate the ATT, the condition that $\{E[Y_{it'}(0)|D_i = 1] - E[Y_{it'}(0)|D_i = 0]\} = 0$, or more intuitively, the treated units would have yielded the same outcomes as the untreated units had the policy not occurred, is presumed.

I will first apply the simplest and most naïve unconditional DID estimator of α_{TT} with two-way fixed effects. Also, as the ETS participating regions literally did not receive the policy treatment at the same year, but rather entered by 3 batches, I will adopt a time-varying form which looks like the following:

$$Y_{it} = \alpha_0 + \alpha_1 ETS_{it} + X_{it}'\beta + \mu_i + \tau_t + \varepsilon_{it} \tag{3}$$

where Y_{it} is the logged term for CO_2 emissions or other outcome variables of interest. $ETS_{it} = 1$ if city i in year t receives the treatment of the ETS pilot scheme [25]. X_{it}' is the vector of controls, in which I include city socioeconomic characteristics and/or city weather conditions. μ_i and τ_t represent city fixed effect and year fixed effect respectively. Additionally, following the argument of Abadie et al. (2017) about adjusting the standard errors, in OLS regressions I will cluster the standard errors at the provincial level. This is due to the policy design that the central government of China assigned the ETS treatment status to a province-level administrative division as a whole, while the data I use is collected at city-level.

As the counterfactual outcomes $E[Y_{it'}(0)]$ in the ATE estimation specified by Equation (2) are impossible to be observed, the DID approach requires the key identifying assumption that the selection bias can be removed by adjusting for some observable differences between treated and control units. More specifically, we are interested in investigating whether there is a parallel trend in the paths of outcome variables before the ETS pilot scheme came into effect between the treatment and control groups. To see if this crucial prerequisite for the DID estimator is satisfied, I adopt the following event study specification:

$$Y_{it} = \alpha_0 + \sum_{j=1}^{J} \alpha_j ETS_{i,t-j} + \sum_{k=0}^{K} \beta_k ETS_{i,t+k} + X_{it}'\omega + \mu_i + \tau_t + \varepsilon_{it} \tag{4}$$

where I include both lag and lead terms of the indicator dummy representing the introduction of ETS between J years before the starting year and K years forward. For example, $ETS_{i,t-3} = 1$ for an ETS participating city in 3 years before the start. I check the existence of parallel trend by inspecting the estimated coefficients α_j on the lagged ETS treatment terms. The basic idea is that if the parallel trend assumption holds, there should be no discrepancy between the outcomes of treatment group and control group and thus the coefficients on lagged ETS treatment terms should not be significantly different from zero. Moreover, we can also view the dynamic changes in the treatment effect by inspecting the evolvement of the estimated coefficients β_k on lead terms for ETS implementation.

5.2 Semiparametric Conditioning Strategy

A concern for the unconditional DID model specification is that it may still induce some biases even though the differences in the observable covariates vector are controlled. Stemming from the intrinsic non-randomized nature

[25] Specifically, in each batch of ETS pilot scheme implementation, the participating regions launched the ETS mostly in the second half of that year. Moreover, for the first batch of the ETS pilots the launch times were mostly after November. Thus I also adopt an alternative coding method that lets the ETS terms in the starting year to be 0.5 and assigns 1 for terms of later years to account for the different lengths of policy exposure between ETS implementing years. The estimation results are almost identical. For the following time-varying DID estimation in this paper I will demonstrate the results with this alternative coding.

of the ETS policy assignment, there are probably some fundamental differences in the characteristics between the treatment group and control group, which indicates that the treatment status is related to the post-treatment outcomes. In this case, the second term in Equation (2) for the selection bias is nonzero and thus the average treatment effect of the ETS pilot scheme on the CO_2 emissions in participating cities cannot be successfully captured by the unconditional DID estimator. Furthermore, due to the difficulty of revealing all possible confounding factors as some of them may even be unobservable or unquantifiable, I therefore apply the conditional DID estimator using propensity score matching (PSM) strategy to deal with this issue.

Based on the idea that the estimation bias in treatment effect induced by the confounding factors can be reduced when the comparison of outcomes is performed using treated and control units with similar characteristics, the PSM strategy uses the propensity score to summarize the multidimensional vector of pre-treatment characteristics into a single-index variable and thus makes the matching feasible. I follow the generalized DID matching estimator proposed by Heckman et al. (1997):

$$\widehat{\alpha_{DID}} = \frac{1}{N_1} \sum_{j \in I_1} \{[Y_{jt'}(1) - Y_{jt^0}(0)] - \sum_{k \in I_0} w_j k [Y_{kt'}(0) - Y_{kt^0}(0)]\} \tag{5}$$

where I_1 denotes the group of the ETS participating cities indexed by j and I_0 denotes the group of nonparticipating cities indexed by k. The weight placed on city k when constructing the counterfactual estimate for treated city j is w_{jk}, which is determined by the similarity in the pre-treatment characteristics vector X_i. As proposed by Caliendo and Kopeinig (2008), the selection is based on the fact that the matching strategy builds on the conditional independence assumption (CIA), which requires that the outcome variable must be independent of treatment conditional on the propensity score. Thus the pre-treatment characteristics vector X_i is supposed to influence simultaneously the decision of assigning the ETS policy treatment and the outcome variable of CO_2 emissions, while remaining unaffected by the introduction of ETS. Following this rule and considering the features of the ETS pilot scheme, I thereby select the variables for matching from the data-available city socioeconomic indicators listed in Panel A of Table 3.

As in most circumstances the matching process based on the selected pre-treatment characteristics generates only one pair of treatment and control groups within the whole data sample, and there are no canonical practice to incorporate matching strategy with the previously used time-varying DID specification. I therefore treat the timing of exposure to the ETS pilot scheme for participating cities to be the same, that is, all participants are assumed to launch the carbon emission trading from the year of 2013 during the matching process. Consequently, the specification of conditional DID estimator I adopt takes the traditional form:

$$Y_{it} = \alpha_0 + \alpha_1 ETSgroup_i \times postETS_t + X_{it}'\beta + \mu_i + \tau_t + \varepsilon_{it} \tag{6}$$

where the time-varying term for ETS is substituted by the interaction term between $ETSgroup_i$ and $postETS_t$. Specifically, $ETSgroup_i = 1$ if city i is an ETS pilot scheme participant and $postETS_t = 1$ in the year of 2013 and hereafter.

Despite that the specification does not capture the staggering policy treatment years across ETS pilot cities, I argue that the drawback in estimating is trivial. This is because the majority of ETS pilots came into effect in 2013 and 2014. Specifically, the first batch of pilot regions launched the ETS mostly at the end of 2013 while the second batch started before mid-2014. Hence, the ETS implementation timings are close and the difference in policy exposure length is minor, thus having very limited impact on the accuracy of conditional DID estimation. In fact, this is also the common method of specifying the DID models in existing literature about the China ETS pilots [26]. To further examine the proposition, I also conduct the the previous time-varying estimation using the matched data and no significant change occurs as shown in Appendix Table 15, indicating the plausibility of the traditional DID specification.

The PSM-DID model in this thesis aims to mitigate the biases in the unconditional DID estimates by matching treatment and control units with similar probabilities (delegated by propensity scores) of being selected to the ETS pilot scheme according to the pre-treatment covariates vector. The estimation under this specification relies on the conditional unconfoundedness assumption that the distribution of the outcome conditional on pre-treatment characteristics vector X_i, is the same among ETS pilot cities and outliers. Another implicit assumption wielded by the causal inference framework in this thesis is the Stable Unit Treatment Value Assumption (SUTVA), which requires that the potential outcomes in one unit are independent of the treatment status of others. Intuitively, SUTVA rules out the circumstances where the ETS treatment effect in the treated cities also interferes the CO_2 emissions in the control group cities. To evaluate the plausibility of these assumptions, I will design and conduct pertinent tests and address the internal validity of my empirical analyses.

Additionally, I impose the common support restriction to improve the quality of propensity score matching. The support of the distribution of the conditioning covariates in the treatment group is required to overlap that in the control group, which ensures that individuals with the same X_i values have a positive probability both of being participants and non-participants. Thereby in the matching process the observations that fail to fall into the common support should be discarded. The most straightforward way to check the overlap in the regions of common support between treatment and control group is to visually inspect the density distribution of the propensity score, so I will plot both the histogram and the kernel density graph of propensity scores for this purpose. Besides, I will evaluate the quality of propensity score matching by comparing the standardised biases before and after matching as suggested by Rosenbaum and Rubin (1985).

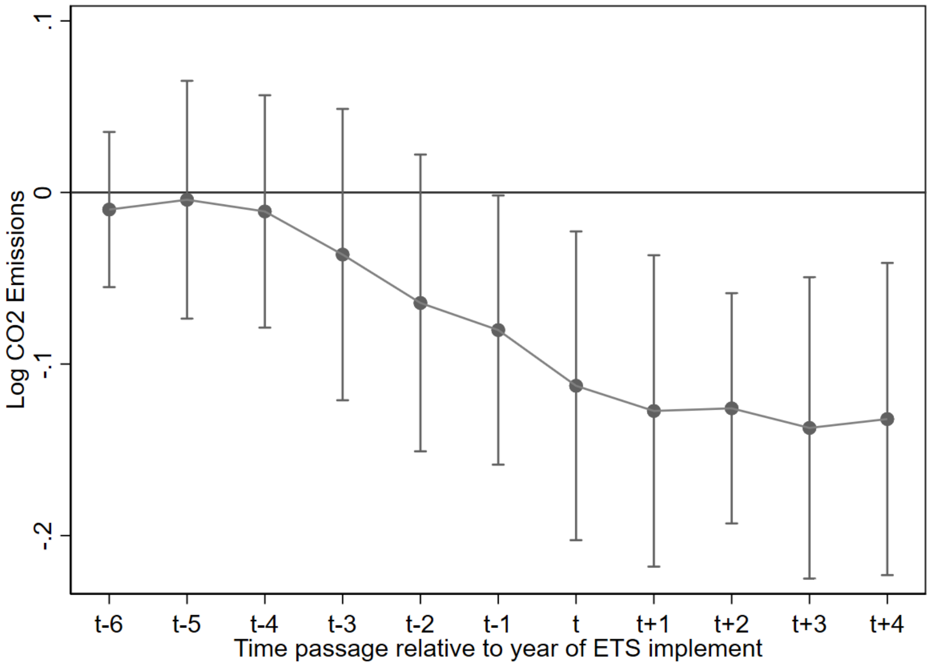
6 Empirical Results

6.1 DID Estimation Without Matching

Under the unconditional DID specification, I first present the event study graphs that motivate the regression analyses that follow. Based on the regression result of the specification in Equation (4), I extract the estimates of the event study coefficients for years before/after the actual implementation of ETS pilot scheme and plot

[26]For example, the papers of Dong et al. (2019), Zhang et al. (2019a), Shen et al. (2020), Feng et al. (2021), etc.

FIGURE 3: EVENT STUDY: THE EFFECT OF ETS ON CITY CO_2 EMISSIONS



Notes : The estimates in Figure 3 are from the event study regression for logged CO_2 emissions specified in Equation (4). Specifically, the time passage spans from 6 years before the ETS implement to 4 years after it. The regression includes city characteristics controls and city and year fixed effects. The standard errors underlying the confidence intervals are clustered at the province level.

them with their 95 percent confidence intervals as shown in Figure 3. Although in most points in time before the implementation there is no significant estimate, the estimated coefficient on t-1 is negative with statistical significance. Moreover, there is a sign of deviation in the paths of CO_2 emissions between treatment group and control group since t-3. This could be due to the anticipation effect that local governments and regulated firms in the pilot cities had made some preparations for the upcoming emission trading policy between the announcement of the work plan of ETS pilot scheme and the official implementation. However, this may also indicate that there exists some foundamental difference between the CO_2 emission pre-trends of treatment group and control group, which impairs the comparability that motivates the adoption of DID approach.

Table 4 demonstrates the unconditional DID regression results. I observe a significant reduction effect of the EST pilot scheme on the CO_2 emissions among the treated cities as shown in Column (1). Next I conduct placebo tests to check the validity of this result. The main idea for the placebo test is to shift the ETS implementing time ahead and redo the DID regression using the subsample before the policy actually starts. If the unconditional DID estimator is properly constructed and yields the estimation for the CO_2 reduction effect attributed solely to the the ETS pilot scheme, the placebo treatment effects are expected not to be statistically distinguishable from zero as the ETS pilot scheme should not have any policy effect before it is introduced. I look into the placebo ETS treatment effects for up to 4 years ahead. Nonetheless, in Columns (2) to (5) all placebo treatment effects are significantly negative. The results are highly against the reliability of the unconditional DID model and suggest

TABLE 4: DIFFERENCE-IN-DIFFERENCES ESTIMATES WITHOUT MATCHING

	Main Regression	Placebo Regressions			
		1 Year Before	2 Years Before	3 Years Before	4 Years Before
	(1)	(2)	(3)	(4)	(5)
ETS	-0.106**	-0.0637**	-0.0790***	-0.0792***	-0.0576***
	[0.002]	[0.003]	[0.001]	[0.000]	[0.001]
	(0.0318)	(0.0196)	(0.0211)	(0.0189)	(0.0161)
Year FE	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Observations	3913	2691	2691	2691	2691
Adjusted R-squared	0.931	0.944	0.944	0.945	0.944

Notes : The entries in Table 4 are the estimated coefficients from the unconditional DID model where the dependent variable is the logged CO_2 emissions in each city/year. Column (1) is for the main regression using the full sample, while Columns (2) to (5) present the coefficient estimates from the placebo regressions using the sample for the years 2003 to 2012 before the ETS pilot scheme was launched. The placebo ETS terms shift up to 4 years ahead of the real policy introduction. In all regressions city characteristics are controlled in which I include population, GDP, GDP per capita, electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and afforestation rate.

Clustered standard errors in parentheses. P-value in square brackets.

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

the existence of confounding factors unaccounted by the model. To cope with the estimation bias originated from model misspecification, I will continue with an alternative approach to identify the treatment effect, namely the Propensity Score Matching Difference-in-Differences (PSM-DID) method, in the following section.

6.2 Propensity Score Matching

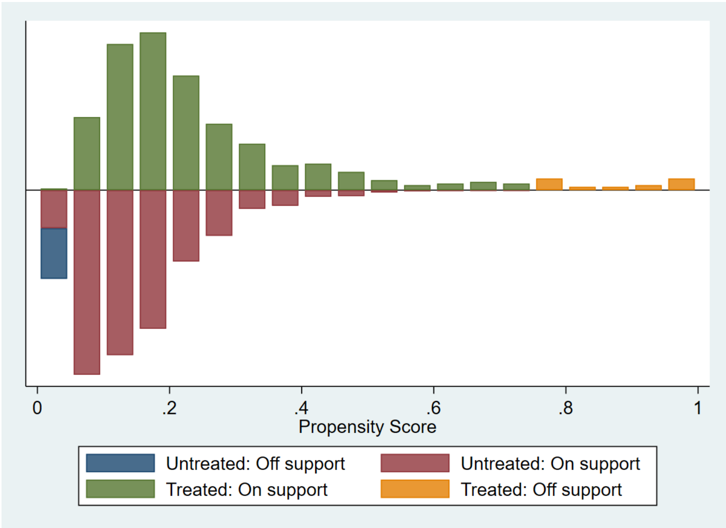
TABLE 5: NEAREST-NEIGHBOR MATCHING RESULT

Treatment Assignment	Off Support	On Support	Total
Untreated	152	2089	2241
Treated	21	429	450
Total	173	2518	2691

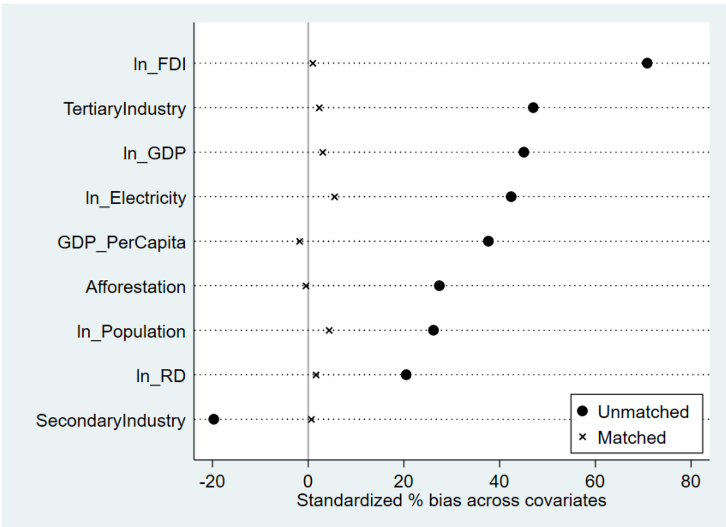
Notes : In the NN matching 4 neighbors are used to calculate the matched outcome. The treatment observations whose pscore is higher than the maximum or less than the minimum pscore of the controls are discarded for the common support restriction. In addition to the nearest neighbor, other control units with identical (tied) pscores are also matched.

The essential ingredient of the PSM-DID analysis is to construct a valid control group that most closely resembles the treatment units. The PSM strategy uses the probability of being exposed to a given policy conditional on the pre-treatment characteristics of the subjects to reduce the dimensionality problem. The identification of the probability of policy exposure, or the propensity score, is based on a logit regression on the chosen pre-treatment covariates. Also, the common support restriction is imposed. Specifically, I will mainly adopt the nonparametric nearest neighbor (NN) matching strategy to construct the counterfactual estimate. Recalling the mathematical specification in Equation (5), the weight w_{jk} equals $\frac{1}{n}$ for the n nearest selected neighbors and equals zero for the rest of the control units. However, as highlighted by Becker and Ichino (2002), a drawback of this matching strategy is that there will be some poor matches due to the fact that the nearest neighbors of some treatment units may have a very different propensity score but they will contribute to the estimation of the treatment effect with equal weight. Therefore alternative matching algorithms are adopted to assess the robustness of the estimates. In Appendix I show that alternative matching strategies will produce similar results.

FIGURE 4: NEAREST-NEIGHBOR MATCHING QUALITY



(A) PROPENSITY SCORE HISTOGRAM BY TREATMENT STATUS

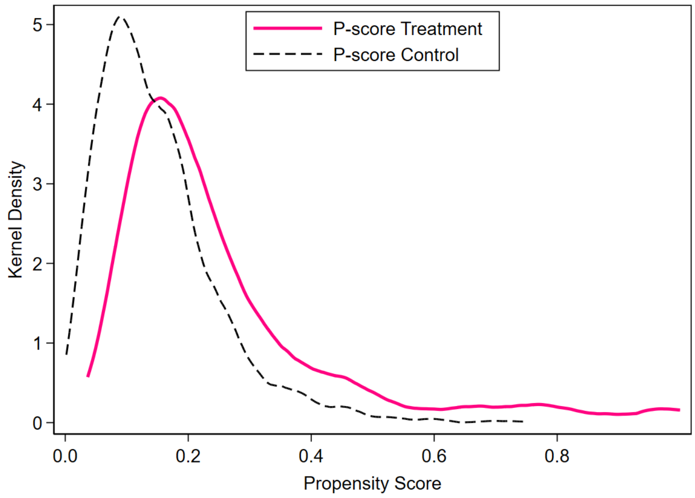


(B) STANDARDISED BIAS BEFORE/AFTER MATCHING

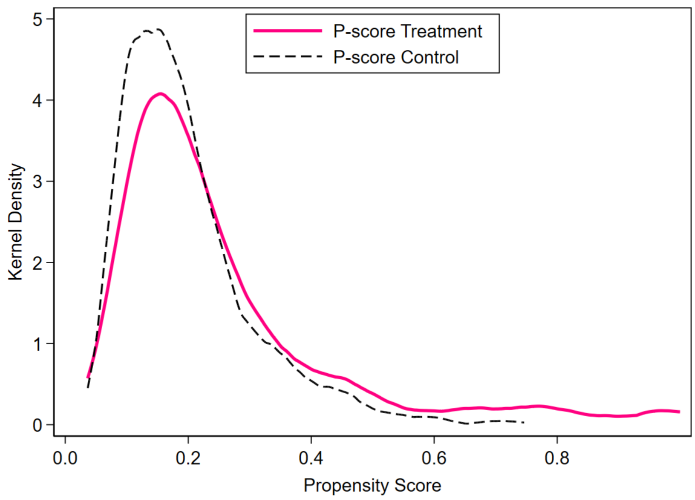
Table 5 summarizes the matching outcomes [27] of the data sample before the initial treatment year of 2013.

[27] For detailed city characteristics variables used for matching, I select logged population, GDP, GDP per capita, total electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and city afforestation rate. To the best of my knowledge and data availability, these socioeconomic indicators should suffice the conditions of influencing simultaneously the ETS treatment assignment decision and the outcome variables while not affected by the treatment.

FIGURE 5: MATCHING QUALITY: KERNEL DISTRIBUTION BEFORE/AFTER NN MATCHING



(A) BEFORE MATCHING



(B) AFTER MATCHING

Among the 2850 observations in the original sample, 2518 are left after matching the propensity scores and sifting for the common support condition. Specifically, 2089 control group observations are matched to 429 observations of ETS pilot scheme participating cities. The number of dropped observations for common support is relatively small, which is not likely to cause the problem that the estimated effect using the remaining individuals becomes nonrepresentative. By simply looking at the joint distributions of the propensity scores in the ETS treatment group and control group as presented in Figure 4a, we can straightforwardly see that the overlap condition is satisfied. To further assess the matching quality, we can see from Figure 4b that the standardised biases for all selected pre-treatment city characteristics, which represent the difference of sample means in the treated and matched control subsamples, are reduced to be around the 5% level after matching. For the more detailed matching outcomes, I present the t-test outcomes in Appendix Figure 9 to show that the differences in the city characteristics vector between groups become indistinguishable from zero. Furthermore, Figure 5 displays the proximity in the kernel density distributions of the propensity scores between treatment and control groups before/after matching, which provides further evidence for the common support after matching. Greenstone and Gayer (2009) propose that one informal method for assessing the internal validity of a quasi-experiment is to test whether the distributions of the observable covariates are balanced across the treatment and control groups. Since the observable pre-treatment characteristics are balanced by the PSM strategy, then it is likely that the unobservables originated from the non-randomized nature of the ETS pilot scheme are also balanced. Overall, the results in this section can to a large extent indicate the success of the matching procedure and its potential to reduce the selection bias. Thereby, I will proceed by conducting the DID analysis based on the matched data sample.

6.3 PSM-DID Estimation Output

TABLE 6: CONDITIONAL DIFFERENCE-IN-DIFFERENCES REGRESSION OUTCOMES

	(1)	(2)	(3)
ETS	-0.0832*	-0.0802*	-0.0834*
	[0.033]	[0.034]	[0.034]
	(0.0372)	(0.0361)	(0.0374)
City Controls	No	Yes	Yes
Weather Controls	No	No	Yes
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	3943	3740	3347
Adjusted R-squared	0.916	0.935	0.935

Notes : The entries in Table 6 are the coefficient estimates from the conditional DID model using the matched data sample. The regressions follow the traditional DID setting that the explaining variable of interest is the interaction between the binary indicators of treatment group and treatment period. Other specifications are identical to the previous unconditional DID regressions. The dependent variable is the logged CO_2 emissions in each city-year. The city characteristics vector contains population, GDP, GDP per capita, electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and afforestation rate. The weather control vector contains the annually averaged temperature, humidity, precipitation and sunshine.

Clustered standard errors in parentheses. P-value in square brackets.

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

With the new data sample composed of matched units from the ETS treatment group and control group obtained through propensity score matching in the preceding sections, I turn to the conditional DID model specification defined in Equation (5) and Equation (6) to estimate the CO_2 emission reduction effect of the ETS pilot scheme. Table 6 demonstrates the estimated treatment effects of the ETS pilot scheme under model specifications using different sets of control variables. From Columns (1) to (3) we can see that all regressions return significant estimates for the ETS treatment effect. The preferred estimate controlling for city characteristics reported in Column (2) suggests that the ETS pilot scheme can reduce the CO_2 emissions in the treatment group cities by 8.02% at a significance level of 0.05. When adding weather controls into the DID regression, the scale of the estimated CO_2 emission reduction effect increases slightly. Compared to the previous unconditional DID estimation results, the estimated CO_2 reducing effect after matching has a smaller size. There are two potential reasons for the difference. On the one hand, the matching process helps eliminate the confounding impacts from the omitted factors which are very likely to be related with the trend of economic and social transform in China over the last decade and thus facilitate the emission reduction. On the other hand, technically the time-varying ETS treatment term under the unconditional DID specification on average implies later implementation timing than in the case of conditional DID which specifies the policy outset for all pilot cities to be the year of 2013. So the time-varying term indicates more intensive exposure to the ETS treatment and thereby probably yields a larger estimation.

Albeit the matching process is proved to be of good quality in previous sections, it should be kept in mind that the propensity matching strategy only serves to reduce rather than to completely exterminate the bias from unobservable confounding factors. Moreover, the PSM-DID estimates represent the average impact of the ETS pilot scheme and provide no evidence for the differentiated policy effect across regions. Therefore, the DID estimation results should be viewed with caution. In pursuance of more concrete causal relationships between the ETS and emission outcomes, I will tackle the concerns about the heterogeneity in policy effect and potential violations to the identifying assumptions in subsequent sections.

6.4 Heterogeneity in Treatment Effect

The unbalanced regional development is typical in China. As a matter of fact, both the inter-provincial and intra-provincial city socioeconomic conditions differentiate considerably. Hence, potential environmental injustice concerns may be cast upon market-based environmental policy like the ETS pilot scheme, as it by design aims to lower the overall CO_2 emission level within the pilot regions and specifies little about the spatial distribution of the emissions. With the "right to pollute" established by the ETS, firms can continue operating and emitting as long as they have purchased the needed emission allowances from the carbon market. Scepticism that pollution hot spots may be generated and especially the disadvantaged areas tend to intake pollution flows are thus thrown at the ETS programs. If the ETS pilot scheme indeed unevenly reduces the CO_2 emissions across different pilot cities, extrapolating the potential average effect of the future nationwide ETS in China from the evaluation of the ETS pilot scheme in this thesis will have a weaker foundation and the external validity of the previous results will be vacillated.

Therefore, beyond the average treatment effect of the ETS pilot scheme, I investigate whether the pilot cities where the residents are more disadvantaged and have lower incomes experience less CO_2 emission reduction than other participating cities so as to address the heterogeneity issue. A Difference-in-Difference-in-Differences approach (triple DID, or DDD) is employed to study this effect. To represent the overall income level of each city, the indicator $LowIncome_j$ is generated and equals 1 for low-income cities. I first define low-income cities as those with the annual wage amount per resident lower than the national median, and thereby cities with higher wage level are assigned to the high income group. In the second setting I use a subsample with more extreme discrepancies that $LowIncome_j$ equals 1 for cities with less than 25 percentile of wage level and $LowIncome_j$ equals 0 for cities with higher than 75 percentile of wage level. I also rerun the previous DID regressions using only the subsamples of low-income cities where the income level is below the national median or 25% percentile to see how the effect estimation would change. The DDD estimator is constructed as:

$$
\begin{aligned}
Y_{ijt} &= \alpha_0 + \alpha_1 ETSgroup_i \times postETS_t \times LowIncome_j \\
&+ \alpha_2 ETSgroup_i \times postETS_t + \alpha_3 ETSgroup_i \times LowIncome_j \\
&+ \alpha_4 postETS_t \times LowIncome_j + X'_{ijt}\beta + \mu_i + \tau_t + \varepsilon_{ijt}
\end{aligned} \tag{7}
$$

where an additional variation in the city income level is exploited compared to the former DID specification in Equation 6. The estimate of coefficient α_1 is of interest as it captures the gap of CO_2 reducing effects between low-income pilot cities and other participants of higher income level.

TABLE 7: RESULTS OF HETEROGENEITY ANALYSIS

	Bottom 50% versus Top 50%		Bottom 25% versus Top 25%	
	DDD	DID	DDD	DID
	(1)	(2)	(3)	(4)
$1(ETSgroup)\times$ $1(postETS)\times$ $1(LowIncome)$	-0.0421		-0.0392	
	[0.374]		[0.525]	
	(0.0464)		(0.0609)	
$1(ETSgroup)\times$ $1(postETS)$	-0.0626	-0.108**	-0.0652	-0.0946**
	[0.159]	[0.002]	[0.289]	[0.002]
	(0.0435)	(0.0308)	(0.0604)	(0.0256)
Year FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Observations	3740	1856	1879	931
Adjusted R-squared	0.935	0.945	0.927	0.943

Notes : The entries in Table 7 are the coefficient estimates from the triple DID model and low-income group DID model. In Column (1) and (3) the explaining variable of interest is the interaction term between the indicators of ETS treatment group, ETS treatment time and low-income group. In Column (2) and (4) the specification follows the previous DID setting and uses the sub sample of low-income group cities. All regressions control for city characteristics.

Clustered standard errors in parentheses. P-value in square brackets.

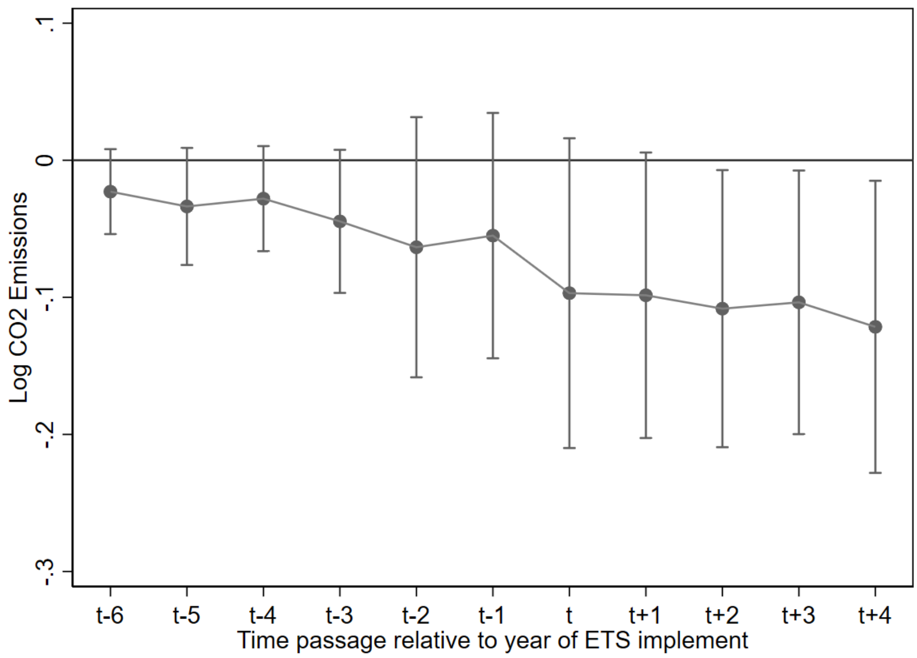
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7 reports the results for the assessment to the heterogeneity in ETS treatment effect. In Columns (1) and (3) for the triple DID specification, there is no statistical significant difference in CO_2 reduction effect between low-income pilot cities and their high-income counterpart within the policy range of ETS pilot scheme, even if I compare the more divergent income groups of bottom 25% and top 25%. Especially, when using the subsamples of low-income cities to do the DID estimation, I find the CO_2 abatement effect in these cities to be larger in scale as shown in Columns (2) and (4). Therefore, no evidence is found to support the existence of environmental injustice induced by heterogeneous effects across cities of different income levels.

6.5 Robustness Checks

6.5.1 Test for Parallel Trend Assumption

FIGURE 6: EVENT STUDY: THE EFFECT OF ETS ON LOCAL CO_2 EMISSIONS AFTER MATCHING



Notes : The estimates in Figure 6 are from the event study regression based on the matched sample and PSM-DID model. Specifically, the time passage spans from 6 years before the ETS implement to 4 years after it. The regression includes city characteristics controls and city and year fixed effects. The standard errors underlying the confidence intervals are clustered at the province level.

Although the growth paths affected by the pre-treatment city characteristics are effectively coordinated via the matching process, the parallel trend assumption for the feasibility of the DID estimator is not guaranteed to hold due to the possibility of unobserved confoundedness. Therefore, as the first-order step for robustness check, I reproduce the event study results using the matched data sample to check this crucial assumption. Slight modifications are made for the event study specification in Equation 4 to incorporate with the traditional DID specification. From Figure 6 we can see clearly that despite the ETS pilot cities are estimated to averagely have lower CO_2 emissions than control group cities before the introduction of ETS, the differences are not statistically significant. Thus the event study is in favor of the scenario that for the matched sample there exits a parallel trend in the CO_2 emission outcomes before the implementation of ETS pilot scheme between treated and control units. The paths of CO_2 emissions start to diverge significantly after the ETS pilots are officially launched. This provides evidence for the legitimacy of the applying the PSM-DID estimator to do causal inference for the treatment effect of ETS pilot scheme.

6.5.2 Placebo Test Using Alternative Treatment Timing and Treated Units

To examine whether the matching process effectively deals with the endogeneity issues that inflicts the previous unconditional DID specification, I redo the placebo tests for the PSM-DID model. In addition to the method of changing the implementation time of ETS pilot scheme in the previous specification, I also conduct a placebo test that changes the composition of the treatment group, based on the straightforward idea that using the non-participating cities to the ETS to form the placebo treatment group should return no significant treatment effect estimates. The falsified treatment group is made up of the cities from the nearby provinces with similar economic conditions [28] to the true ETS participating provinces and municipalities. I use this fictitious treatment group to go through the PSM-DID matching and estimating process all over again. If the matching strategy mitigate the impact of confounding factors, I expect to find none of the placebo ETS effect estimates to be significant.

TABLE 8: PLACEBO TESTS FOR PSM-DID MODEL

	Start in 2012 (1)	Start in 2011 (2)	Start in 2010 (3)	Start in 2009 (4)	Start in 2008 (5)	Falsified Group (6)
ETS	-0.0329	-0.0383	-0.0369	-0.0347	-0.0370	0.0425
	[0.254]	[0.286]	[0.239]	[0.190]	[0.140]	[0.271]
	(0.0283)	(0.0352)	(0.0307)	(0.0259)	(0.0244)	(0.0378)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2518	2518	2518	2518	2518	3857
Adjusted R-squared	0.948	0.949	0.949	0.947	0.947	0.931

Notes : The entries in the first 5 columns of Table 8 are the coefficient estimates from the placebo tests that shift the actual ETS implementing times up to 5 years ahead and regress with the matched data sample for the years 2003 to 2012 before the ETS pilot scheme was launched. The results for placebo test that changes treatment group composition are in Column (6). In all placebo regressions city characteristics are controlled.

Clustered standard errors in parentheses. P-value in square brackets.

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

The placebo regression results are shown in Table 8. From Columns (1) to (5), the placebo regressions in which the ETS implementation is shifted earlier for up to 5 years are all returning insignificant estimates of CO_2 reducing effect. Column (6) reports the estimating result for the placebo treatment group. Again no significant treatment effect can be observed. Thus, combined with the propensity score matching strategy the conditional DID model remarkably reduce the impact of unobserved confounders and strengthen the robustness of the estimating results.

6.5.3 Check for the Uniqueness of the ETS Treatment Effect

In the assessment for the CO_2 reduction effect of ETS pilot scheme, other environmental policies that are issued by the China central government during the research time range of this thesis may lead to underestimation or exaggeration to the ETS policy impact. Among the relevant policies the most suspicious one is the Low-Carbon

[28] To put in detail, I construct the set of treatment units to be the cities in Hebei Province, Liaoning Province, Zhejiang Province, Jiangxi Province, Hunan Province, Guangxi Province, Hainan Province and Sichuan Province.

TABLE 9: DID ESTIMATION INCLUDING LCC POLICY

	(1)	(2)	(3)
ETS	-0.0693	-0.0740*	-0.0784*
	[0.087]	[0.049]	[0.046]
	(0.0392)	(0.0360)	(0.0376)
LCC	-0.0317	-0.0154	-0.0125
	[0.387]	[0.561]	[0.648]
	(0.0361)	(0.0262)	(0.0270)
City Controls	No	Yes	Yes
Weather Controls	No	No	Yes
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	3943	3740	3347
Adjusted R-squared	0.916	0.935	0.935

Notes : The entries in Table 9 are the coefficient estimates from the DID specification in Equation (8). The matched data sample is used for regressions following the same setting as the previous PSM-DID models. The dependent variable is the logged CO_2 emissions in each city-year. The city characteristics vector contains population, GDP, GDP per capita, electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and afforestation rate. The weather control vector contains the annually averaged temperature, humidity, precipitation and sunshine.

Clustered standard errors in parentheses. P-value in square brackets.

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

City (LCC) project, a CAC-type policy that targeted exactly at CO_2 emissions. In 2010, the China NDRC (National Development and Reform Commission) issued the *Notice on Carrying Out Pilots of Low-Carbon Provinces and Cities*, calling for dozens of low-carbon city pilots to be launched around the country. In July 2010, the LCC project is launched in 5 provinces, 2 centrally-administered municipalities and 5 cities. In the end of 2012, the second batch of LCC project incorporated 1 province, 2 centrally-administered municipalities and 24 cities [29]. Within the program, local governments carry out reforms such as promoting green buildings and higher energy-saving criteria in urban infrastructure and facilities. The central government also makes funding available to support the LCC complying cities to invest in energy efficiency and renewable energy capacity technology, either through grants or preferential financing from the China Development Bank and other policy banks. Furthermore, achievements in complying with the LCC targets are included in the metrics of evaluating the performance of municipal officials. According to China's First Biennial Update Report (December 2016) to the UNFCCC, CO_2 emissions per unit of GDP in LCC complying regions fell 19.4% from 2010 to 2014, faster than the national average [30].

Most importantly, as shown in Appendix Figure 10 of the spatial distribution map for the LCC regions, the vast majority of the ETS pilot cities [31] had been complying with the LCC project before launching the ETS, which may cause the overstatement of the previous estimation of ETS treatment effect. And since the LCC policy was

[29] https://chineseclimatepolicy.energypolicy.columbia.edu/en/low-carbon-cities

[30] https://unfccc.int/sites/default/files/resource/3_China_FSV_Presentation.pdf

[31] Except for a few cities in Fujian province.

delivered by the central government and implemented in a specific group of cities, the two-way fixed effects may fail to account for it. Thereupon, the risk that previous estimate for the ETS treatment effect is mixed up with the policy effect of LCC may undermine the credibility of causal inference. To probe the influence of LCC on my previous analysis for the ETS pilot scheme, I modify the DID estimation in Equation (6) to incorporate the LCC policy:

$$Y_{it} = \alpha_0 + \alpha_1 ETSgroup_i \times postETS_t + \alpha_2 LCCgroup_i \times postLCC_t + X_{it}'\beta + \mu_i + \tau_t + \varepsilon_{it} \tag{8}$$

where the DID term for LCC project is added. Specifically, $LCCgroup_i = 1$ if city i is an LCC participant and $postLCC_t = 1$ in the year of 2010 and hereafter. I am particularly interested in checking whether there will be a significant drop in the estimate of the coefficient α_1 after introducing the LCC policy in the model.

Table 9 shows the outcomes for the DID estimation giving consideration to the LCC policy. From Column (2) I find the preferred estimate of the ETS treatment effect to be -7.40% at the 95% confidence level. Indeed there is a decrease in the estimated CO_2 reduction effect after accounting for the LCC policy, nevertheless, this small-sized drop in the effect of ETS is rather feeble to foil the previous results. More so, in all regressions of Table 9 the estimate of the LCC emission reduction effect is not distinguishable from zero. Therefore, it is reasonable to believe that the previously estimated effect of ETS is not very sensitive to the impacts of other CO_2 emission regulating policies. Moreover, the comparative analysis between ETS pilot scheme and LCC project in this part also sheds some light on the superiority of the market-based environmental policies over the CAC policies in regard to the efficacy of reducing the targeted emissions.

6.5.4 Cope with Serial Correlation in Treatment Variable: Data Aggregation Technique

Apart from testing the identifying assumptions of the DID model, I also consider the potential estimation bias induced by the construction of the DID estimator used in the thesis. Specifically, when applying the Difference-in-Differences technique the employed data mostly contain fairly long time periods and the dependent variables are often positively serially correlated. Besides, an intrinsic aspect of the DID model is that in most cases the policy treatment status of a unit remains unchangeable for a long period. Therefore, as highlighted by Bertrand et al. (2004), this characteristic of DID model may give rise to biases in estimating the standard error around the treatment effect and result in dramatically higher rejection rates of the null hypothesis of no effect. Correspondingly, they propose several techniques to deal with this serial correlation problem, one of which is to remove the time-series dimension by aggregating the data into pre-treatment period and post-treatment period.

The panel data used in this thesis cover a total of 284 cities, which make the sample size sufficiently large for the data aggregation technique to function well. I thus average the data before and after the implementation of ETS pilot scheme and redo the regression in Equation (6) with a 2-period panel. The results for the aggregated data regression is reported in Table 10. We can see that after ruling out the time series components in the data and correcting for standard error estimation, in Column (1) where no control variable is included the estimate becomes insignificant. Nonetheless, when applying the ideal specification of controlling city characteristics, the estimated CO_2 reduction effect of ETS pilot scheme is still statistically significant, and the scale (-8.16%) is highly

TABLE 10: REGRESSION OUTCOME USING AGGREGATED DATA

	(1)	(2)	(3)
ETS	-0.0349	-0.0816*	-0.0917**
	[0.235]	[0.028]	[0.009]
	(0.0287)	(0.0353)	(0.0327)
City Controls	No	Yes	Yes
Weather Controls	No	No	Yes
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	541	530	523
Adjusted R-squared	0.900	0.926	0.932

Notes : The entries in Table 10 are the coefficient estimates from the PSM-DID model under the data aggregation setting. Within the matched dataset, the outcome variable of CO_2 emissions, the city characteristics controls as well as the weather controls are all aggregated before/after the year of 2013. Regressions are then conducted with the 2-period panel.

Clustered standard errors in parentheses. P-value in square brackets.

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

consistent with the previous result. Interestingly, the estimation result using both city characteristics controls and weather controls now suggests the CO_2 emission reduction to be higher than 9%. The aggregated regression results to a substantial degree dispel the concerns about the over-rejection of the previous DID estimation with long panel and reassure the reliability of the causal inference in the thesis.

6.5.5 Test for the Stability of PSM-DID Estimates

For the consistency of estimating the CO_2 emission reduction effect of the ETS pilot scheme, the Stable Unit Treatment Value Assumption (SUTVA) is desired. It requires that the ETS treatment received by the participating cities does not affect the CO_2 emission outcomes in other cities of the control group. Considering the possibility that the regulated firms or facilities in pilot cities may have the attempt to relocate in order to dodge the regulation of ETS, the counterfactual estimates in this paper may be biased by the carbon leakage. Hence the risk of overstating the ETS treatment effect arises in my analysis. To cope with such concern, I evaluate the stability of the treatment effect in an indirect manner. Following the reasoning of Fowlie et al. (2012), two hypotheses about the channels that the stability of the ETS treatment effect degenerates are proposed:

Hypothesis 1. The ETS pilot scheme induces the regulated firms to shift their production and CO_2 emissions to the nearby non-participating cities

Hypothesis 2. The ETS pilot scheme induces the regulated firms to shift their production and CO_2 emissions to the cities with less stringently environmental regulations

To test these two hypotheses for SUTVA, I separate the observations of the control group into different subsets to identify the sample average treatment effect. Specifically, for the first hypothesis about the carbon leakage to nearby non-ETS cities, I exclude the cities in the adjacent provinces to ETS pilot regions from the control group. If Hypothesis 1 is true, then I would expect to see a shrinkage in ETS treatment effect estimated from this subset of data. For Hypothesis 2 I define the less stringently regulated status of a city as not being registered in the Low-Carbon City project. To the best of my knowledge, I am not aware of other centrally issued environmental policies regulating CO_2 emissions that could induce variance to the local emission regulation stringency in China over the recent decade. So I deem that it is plausible to use the LCC complying status as the identifier. Moreover, to maintain a larger size for the remaining sample, the LCC participating cities are dropped from the control group. The DID regression with only less stringently regulated cities in the control group is supposed to produce higher estimates for the ETS treatment effect if Hypothesis 2 holds.

TABLE 11: DID REGRESSIONS FOR STABILITY TEST

	Whole Data Sample (1)	No Adjacent Regions (2)	No Stringent Regions (3)
ETS	-0.0802*	-0.0869*	-0.0781*
	[0.034]	[0.035]	[0.049]
	(0.0361)	(0.0387)	(0.0377)
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	3740	2341	3006
Adjusted R-squared	0.935	0.922	0.935

Notes : The entries in Table 11 are the coefficient estimates from both the whole data sample and the subsamples. In Column (2) the observations for adjacent provinces (Hebei, Liaoning, Zhejiang, Henan, Guangxi, Hunan, Sichuan) are dicarded, while in Column (3) the LCC cities from the control group are dropped. Other regression specification is the same as before. The city characteristics are controlled.

Clustered standard errors in parentheses. P-value in square brackets.

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

Table 11 shows the comparison between the DID estimation using original data sample and subsamples. When excluding adjacent regions from the control group, the estimate of the CO_2 reduction effect does not decline, but rather increase a little bit to be -8.69% as observed in Column (2). Therefore I can reject Hypothesis 1 that the effect is exaggerated by carbon leakage to nearby non-ETS regions. Similarly, Column (3) reports a slightly lower estimation of CO_2 abatement effect of -7.81%, which is contradictory to the expectation of a higher estimate from Hypothesis 2. Hence there is no evidence that the ETS regulated firms would move to and pollute in cities with lax environmental policies, either. As both hypotheses are rejected, I argue that the ETS treatment effect estimation is not significantly impacted by the composition of control group. The PSM-DID estimation results exhibit an ideal level of stability and they are in favor of the scenario that the SUTVA is not violated.

7 Other Impacts of the ETS Pilot Scheme

Apart from the main results for the CO_2 emission reduction effect analysis, I further look into the synergistic impacts of the China ETS pilot scheme on other relevant outcome variables in this section. The following contents will be devoted to the investigation about two aspects of interest: the local air quality and the local economy.

7.1 The Local Air Quality

TABLE 12: PM2.5 AS THE OUTCOME VARIABLE

	(1)	(2)	(3)	(4)	(5)	(6)
ETS	-0.135**	-0.115*	-0.118*	-0.131**	-0.110**	-0.113**
	[0.008]	[0.016]	[0.018]	[0.002]	[0.007]	[0.007]
	(0.0471)	(0.0450)	(0.0473)	(0.0383)	(0.0377)	(0.0388)
LCC				0.00926	0.0109	0.0118
				[0.798]	[0.771]	[0.761]
				(0.0359)	(0.0371)	(0.0383)
City Controls	No	Yes	Yes	No	Yes	Yes
Weather Controls	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3930	3732	3347	3930	3732	3347
Adjusted R-squared	0.394	0.420	0.382	0.394	0.420	0.382

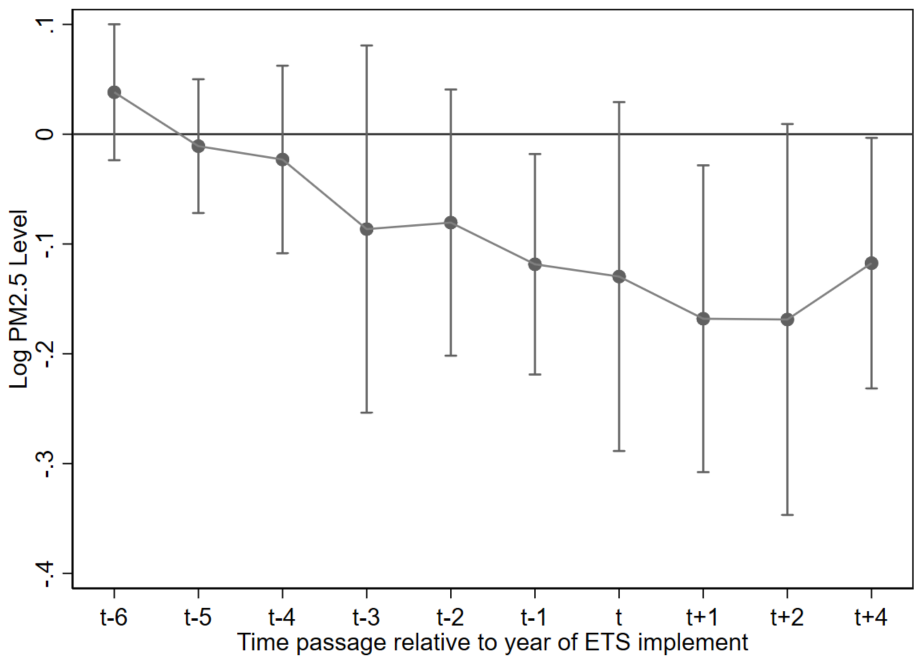
Notes : The entries in Table 12 are the coefficient estimates from the conditional DID model using the matched data sample. The regressions follow the traditional DID setting that the explaining variable of interest is the interaction between the binary indicators of treatment group and treatment period. In Columns (1) to (3) the model specification follows the Equation (6). In Columns (4) to (6) the model specification follows the Equation (8) where the DID term for LCC policy is added. The outcome variable is substituted to be the logged PM2.5 level in each city-year. The city characteristics vector contains population, GDP, GDP per capita, electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and afforestation rate. The weather control vector contains the annually averaged temperature, humidity, precipitation and sunshine.

Clustered standard errors in parentheses. P-value in square brackets.

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

Despite the fact that China ETS pilot scheme is essentially a CO_2 emission regulating program and does not directly target at the local criteria air pollutants, it is reasonable to consider the scenario that the regulated firms' efforts to comply with the CO_2 reduction goals might also bring about the decline in emissions of other air pollutants and thus improve the local air quality. This could be achieved through the channels such as the usage of cleaner fuel or the innovation on production technology. As the air quality is closely linked to the health of local residents, evaluating the synergistic effect of ETS on local air pollutants can help to form a more comprehensive perspective on the effect of the ETS pilot scheme. Thereby, I use the city ground-level PM2.5 measurements as the dependent variable and repeat the previous DID regressions. Table 12 reports the results for the impact of ETS on PM2.5. All the estimates are significantly negative and suggest that ETS can decrease the PM2.5 level in the participating cities. From Column (2) where city characteristics are controlled, we can see that the PM2.5 mitigation effect is -11.5%. And in Column (3) where the weather conditions are also accounted for, the estimated effect is -11.8%. Additionally, I control for the impact of LCC policy in Columns (4) to (6) and observe slightly decreases in the scale of the estimated PM2.5 mitigation effect, nevertheless, the statistically significance for all estimates increases. The estimated effect for LCC policy on PM2.5 levels is minor and not distinguishable from zero.

FIGURE 7: EVENT STUDY: THE EFFECT OF ETS ON LOCAL PM2.5 LEVELS



Notes : The estimates in Figure 7 are from the event study regression based on the matched sample and PSM-DID model.The outcome variable is substituted to be the logged PM2.5 level in each city-year. Specifically, the time passage spans from 6 years before the ETS implement to 4 years after it. The regression includes city characteristics controls, weather controls, and city&year fixed effects. The standard errors underlying the confidence intervals are clustered at the province level.

To further check the credibility of attributing the drop in PM2.5 levels to the ETS pilot scheme, I reproduce the event study graphs and the placebo tests of alternative ETS implementation years. However, from the event study results displayed in Figure 7, the gap between PM2.5 levels in the treatment group and control group is significantly different from zero a year before the introduction of ETS pilot scheme [32]. The existence of latent confounding factors is further validated by the placebo tests presented by Table 13. In Panel A where the the LCC policy is not included, the placebo PM2.5 mitigation effect is significant in 2012 and 2011 before ETS pilot scheme was launched. Also recalling the previous surface PM2.5 concentration level trajectory plot from Figure 2b, we can visually perceive the significant drop in the year 2012. Even when I account for the impact of the LCC policy in Panel B, the placebo treatment effect on local PM25 levels is still significant in 2012 and 2011, although the scale and statistical significance of the estimates decrease comparing to the previous placebo regressions where the LCC policy is not accounted. This result shows that considering only the LCC policy fails to reflect the whole picture of the underlying contributors to the pre-ETS declining trend of local surface PM2.5 level.

Therefore, the empirical analysis using PM2.5 levels as the outcome variable cannot facilitate a concrete causal inference and there remains ambiguity in the estimated mitigation effect of the ETS pilot scheme. Some important

[32] The event study regression specification for Figure 7 does not include the DID term for the LCC policy. Nevertheless, the graph output is visually identical and the coefficient at t-1 is still significant negative when adding it into the regression.

TABLE 13: PLACEBO TESTS FOR PM2.5 MITIGATION EFFECT

	Start in 2012 (1)	Start in 2011 (2)	Start in 2010 (3)	Start in 2009 (4)	Start in 2008 (5)
Panel A. Placebo tests not controlling LCC policy					
ETS	-0.0916**	-0.0776**	-0.0857	-0.0678	-0.0534
	[0.005]	[0.006]	[0.059]	[0.120]	[0.137]
	(0.0302)	(0.0263)	(0.0437)	(0.0424)	(0.0349)
Adjusted R-squared	0.355	0.357	0.361	0.358	0.355
Panel B. Placebo tests controlling LCC policy					
ETS	-0.0804*	-0.0684*	-0.0834	-0.0610	-0.0458
	[0.016]	[0.024]	[0.096]	[0.172]	[0.202]
	(0.0313)	(0.0287)	(0.0485)	(0.0435)	(0.0351)
Adjusted R-squared	0.356	0.357	0.361	0.358	0.356
Year FE	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Observations	2376	2376	2376	2376	2376

Notes : The entries in Table 13 are the coefficient estimates from the placebo tests for PM2.5 mitigation effect that shift the actual ETS implementing times up to 5 years ahead and regress with the matched data sample for the years 2003 to 2012 before the ETS pilot scheme was launched. In all placebo regressions city characteristics and weather conditions are controlled.

Clustered standard errors in parentheses. P-value in square brackets.

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

factors may be omitted in the estimation. For instance, since the year of 2008, the U.S. Embassy in Beijing has started posting data from an air-quality monitor. In 2011 the discrepancy between the official air quality report in Beijing and the data posted by the U.S. Embassy aroused great concern of the society. The public as well as the government has paid more attention to the air quality issues. In 2013, the Chinese government began publishing the air quality index (AQI) and monitoring the real time PM2.5 which was not listed in the measurement of air quality before [33]. Future research effort may be devoted to capturing the influence of the latent confounding factors and extracting a cleaner estimate for the PM2.5 mitigation effect of ETS pilot scheme.

7.2 The Local Economy

The thesis so far has concentrated on the investigation about the environmental improvements delivered by the China ETS pilot scheme. Nevertheless, it should be kept in mind that the benefit of the ETS may also be accompanied with economic cost. The existing research about the cost induced by the environmental regulation mainly focused on the shock to the employment and production. I will also inspect the regulation cost from this angle.

[33] https://www.sciencemag.org/news/2018/04/rooftop-sensors-us-embassies-are-warning-world-about-crazy-bad-air-pollution

TABLE 14: ESTIMATION FOR THE IMPACT OF ETS ON LOCAL ECONOMY

	Power Sector Employment			Industrial Output		
	DDD	DDD	DID	DDD	DDD	DID
	(1)	(2)	(3)	(4)	(5)	(6)
$1(ETSgroup)\times$	-0.0944			0.0634		
$1(postETS)\times$	[0.128]			[0.359]		
$1(LowIncome)$	(0.0603)			(0.0680)		
$1(ETSgroup)\times$		0.049			-0.120	
$1(postETS)\times$		[0.390]			[0.052]	
$1(IndustrialCity)$		(0.0569)			(0.0592)	
$1(ETSgroup)\times$	-0.0713	-0.0641	-0.0265	-0.0442	0.0898	-0.0133
$1(postETS)$	[0.208]	[0.392]	[0.566]	[0.385]	[0.111]	[0.775]
	(0.0554)	(0.0738)	(0.0457)	(0.0501)	(0.0546)	(0.0461)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3739	3739	3739	3734	3734	3734
Adjusted R-squared	0.075	0.056	0.054	0.846	0.846	0.845

Notes : The entries in Table 14 are the coefficient estimates from the DID models and the triple DID models for city heterogeneity in income or dependency on secondary industry. In Column (1) and (4) the explaining variable of interest is the interaction term between the indicators of ETS treatment group, ETS treatment time and low-income group. In Column (2) and (5) the explaining variable of interest is the interaction term between the indicators of ETS treatment group, ETS treatment time and industrial-city group. In Column (3) and (6) the specification follows the previous DID setting and uses the sub sample of low-income group cities. All regressions control for city characteristics.

Clustered standard errors in parentheses. P-value in square brackets.

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

Notably the China ETS pilot scheme regulates mainly the power sector (including combined heat and power, as well as captive power plants of other sectors). Thereby I will check if there is any detriment to the jobs in this sector using the city-level power sector employment data. As for the impact on production, I exploit the city-level industrial output data to evaluate the ETS policy shock.

Based on the previously matched data, I first adopt the DID model specification in Equation(6) to estimate the average treatment effect of ETS on these two outcome variables. The corresponding event study plots are also provided to demonstrate the dynamics of the effect. Then I adopt the previous heterogeneity analysis specification in Equation (7) to see if low-income cities in the ETS treated group are more heavily impacted by the introduction of ETS pilot scheme. Additionally, I consider the potential heterogeneity in cities of different local economy structures. Out of the concern that cities whose economy depends more on the secondary sector may be more sensitive to the policy shock induced by the ETS pilot scheme, I divide the cities into 2 groups according to their GDP's dependency on secondary sector output [34] denoted by the dummy variable $IndustrialCity_j$. Combining this identifier with the treatment status of ETS, I construct a triple DID model following the specification in Equation

[34] Specificaly, cities whose ratio of secondary sector output in GDP is higher than that of primary industry and tertiary industry are thus classified in "industrial city" group.

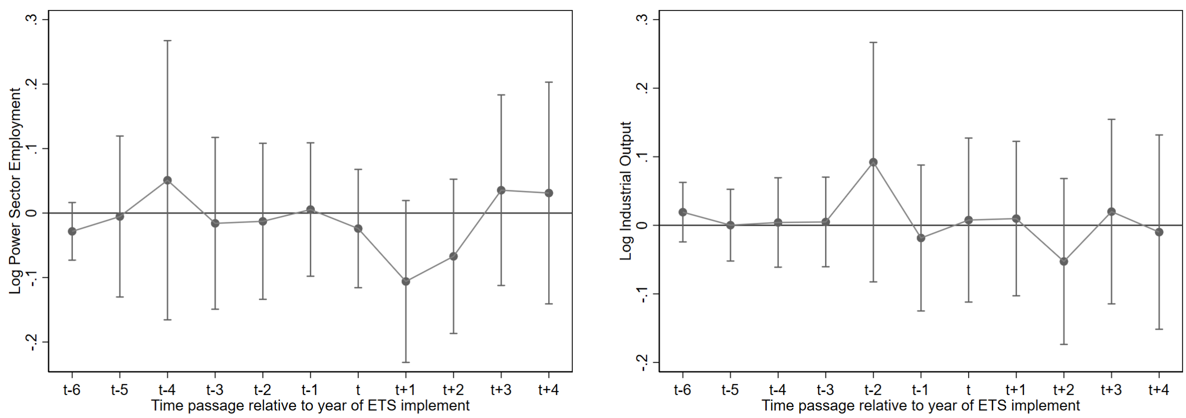
(7): The DDD estimator is constructed as:

$$
\begin{aligned}
Y_{ijt} &= \alpha_0 + \alpha_1 ETSgroup_i \times postETS_t \times IndustrialCity_j \\
&+ \alpha_2 ETSgroup_i \times postETS_t + \alpha_3 ETSgroup_i \times IndustrialCity_j \\
&+ \alpha_4 postETS_t \times IndustrialCity_j + X'_{ijt}\beta + \mu_i + \tau_t + \varepsilon_{ijt}
\end{aligned} \tag{9}
$$

where Y_{ijt} represents the power sector employment or the industrial production. The estimate of coefficient α_1 is of interest as it reflects the heterogeneous impact of the ETS on the pilot cities with high reliance on the secondary sector.

The event study graphs in Figure 8 indicate no significant inter-group divergence in the two outcome variables before and after the introduction of ETS pilot scheme. Similarly, the DID regression results reported in Columns (3) and (6) of Table 14 also suggest no statistically significant impact from ETS on power sector employment or industrial output. Notwithstanding the insignificance in the estimated average impact, it is noteworthy that in the event study plot both the power sector employment and industrial output exhibit a certain downward trend during the initial years of ETS implementation before phasing out in later years. To identify whether or not this stems from the heterogeneity across cities, I then look at the triple DID estimations demonstrated in Columns (1) and (4). Again I find no sign of heterogeneous ETS policy shock to the local economy between low-income pilot cities and those with higher income levels. When investigating the heterogeneity between industrial cities and non-industrial cities, the triple DID regression results reported in Columns (2) and (5) still suggest no significant differences. Nevertheless, it is noteworthy that with respect to the industrial output, the industrial cities within the ETS treatment group would suffer from a 12% higher reduction than others at a statistical significance level

FIGURE 8: EVENT STUDY: THE EFFECT OF ETS ON LOCAL ECONOMY



(A) IMPACT ON POWER SECTOR EMPLOYMENT

(B) IMPACT ON INDUSTRIAL OUTPUT

Notes : The estimates in Figure 8 are from the event study regression based on the matched sample and PSM-DID model. The outcome variable is substituted to be the logged Power Sector Employment and Industrial Output level in each city-year. Due to the missing industrial output data in 2017, I fill the blank using the 2016 output level. Specifically, the time passage spans from 6 years before the ETS implement to 4 years after it. The regression includes city characteristics controls and city&year fixed effects. The standard errors underlying the confidence intervals are clustered at the province level.

pressing on towards 0.05. This finding echos the prior assumption that cities with higher reliance on the secondary sector in their economy are more vulnerable to the ETS policy shock. In general, the aforementioned estimation results are inclined to the judgement that the ETS pilot scheme would not deliver salient damage to the jobs of the main regulated sector nor the industrial production in treated cities.

8 Concluding Remarks

This thesis aims at evaluating the environmental effect of the China Carbon Emission Trading System Pilot Scheme with the city-level panel data from 2003 to 2017. By exploiting the quasi-natural experiment trait of the ETS pilot scheme, I apply the Difference-in-Differences estimator as well as the Propensity Score Matching technique to extract its CO_2 emission reduction effect. The analysis in this thesis contributes to the existing literature that assess the performance of the China ETS pilot scheme in the following three aspects. Firstly the analysis in this thesis is based on the city-level data, which can capture more variations induced by the ETS pilot scheme and improve the precision of estimating its environmental effects. Secondly, the thesis proposes various sources of potential confounding factors to the treatment effect of the ETS pilot scheme and constructs a series of tests to deal with them, thus strengthening the validation and stability of the CO_2 reduction effect estimations. Thirdly, the thesis discusses indirect impacts of the ETS pilot scheme on the local economy and air quality, which offers a more comprehensive view on the benefits and costs of the policy.

The main finding of the thesis is that the ETS pilot scheme significantly reduced the CO_2 emissions by approximately 8% in the pilot cities, on average, relative to the non-participating cities. This estimate is in line with the majority of the existing literature which recognize the CO_2 reducing effect of the ETS pilot scheme, although the scale of the estimated effect in my thesis is more modest. The results are based on the DID estimation combined with semiparametric matching strategy, which performs better than the single DID specification as highlighted in the main body of the thesis. The matching is of good quality and no significant placebo ETS treatment effect is detected under this setting, suggesting that it could mitigate the selection bias stemming from the non-randomized treatment assignment. The PSM-DID estimating outcomes further endure a bunch of robustness checks, through which I rule out the confounding effect of LCC policy that also regulates CO_2 emissions and reject the hypotheses about carbon leakage to regions of the control group. Additionally, I find no evidence of less CO_2 emission reduction in low-income pilot cities. Furthermore, the assessment of the indirect impacts the ETS pilot scheme in general indicates no apparent negative shock to the industrial production or power sector employment, though the industrial output in industrial cities tends to be affected more than that in other cities. With regard to the local air quality measured by the PM2.5 concentration level, the estimated impact from ETS pilot scheme remains ambiguous due to the existence of other contributing factors. Turning back to the main study of this thesis about the CO_2 emission reduction effect, since I have found little evidence for the heterogeneity or instability in the ETS treatment effect, it is plausible to extrapolate the performance of the ETS pilot scheme for anticipating the effectiveness of the futuristic national ETS in China. On the whole, this thesis concludes that the effect of the China ETS pilot scheme is progressive, and implies a optimistic perspective towards what could be delivered by the China's forthcoming national ETS that will be more extensive in covered sectors and more stringently regulated. We can expect that the related efforts made by China in carbon neutrality can make a difference in coping with the climate change.

There are several limitations in the thesis. Firstly, the China CO_2 emission data is originally measured at the county level. However, due to the fact that the official county-level statistical yearbooks are poorly organized and the deficiencies in some important indicators are severe, I was not able to gather a desirable socioeconomic dataset at county level. Therefore I made the compromise by aggregating the county-level CO_2 emissions into city-level statistics for compatibility with the available city characteristics data. This inevitably causes certain losses in the amplitude of data and thus the variations the thesis can exploit. Secondly, for the estimates of PM2.5 mitigation effect from the ETS, the underlying parallel trend assumption is violated. Thus this thesis cannot distinguish the impact of the ETS pilot scheme from other environmental policies or events that swerve the trajectory of local PM2.5 concentration levels. Another drawback lies in the cost analysis for the ETS pilot scheme. The readers should note that the thesis only provides a crude estimation for the potential impact on local industrial production based on the available city-level data. The variations in the aggregate industrial output may not very well capture the more subtle consequences to the performance of the regulated sectors, which calls for caution in interpreting the relevant results. Future research could pursue a more precise estimate for ETS treatment effect by constructing an ideal county-level dataset, since county is the most basic governmental unit in China and can better capture the regional heterogeneity than cities or provinces. Moreover, to draw more concrete causal inferences for the impact of ETS pilot scheme on local air quality and local economy, additional details about the regulated sectors and the concerned employees should be exploited and other potential confounding factors should also be accounted for. Future research could make some firm-level or individual-level investigations within the relevant sectors. Finally, the China ETS exhibits some different features in its design from the existing carbon trading programs elsewhere in the world. For instance, unlike the EU ETS which sets an exogenous cap on the carbon emissions, the China ETS relies on a tradeable performance standard through which a firm's emission allowance is endogenous to its production level. A potential topic for future research could be the assessment for the unconventional mechanisms of China ETS.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1):1–19.

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3):399–424.

Bayer, P. and Aklin, M. (2020). The european union emissions trading system reduced co2 emissions despite low prices. *Proceedings of the National Academy of Sciences*, 117(16):8804–8812.

Becker, S. O. and Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The stata journal*, 2(4):358–377.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1):31–72.

Cao, L., Tang, Y., Cai, B., Wu, P., Zhang, Y., Zhang, F., Xin, B., Lv, C., Chen, K., and Fang, K. (2021). Was it better or worse? simulating the environmental and health impacts of emissions trading scheme in hubei province, china. *Energy*, 217:119427.

Chen, J., Gao, M., Cheng, S., Hou, W., Song, M., Liu, X., Liu, Y., and Shan, Y. (2020). County-level co 2 emissions and sequestration in china during 1997–2017. *Scientific Data*, 7(1):1–12.

Corburn, J. (2001). Emissions trading and environmental justice: distributive fairness and the usa's acid rain programme. *Environmental Conservation*, pages 323–332.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.

Deschênes, O., Greenstone, M., and Shapiro, J. S. (2017). Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review*, 107(10):2958–89.

Dong, F., Dai, Y., Zhang, S., Zhang, X., and Long, R. (2019). Can a carbon emission trading scheme generate the porter effect? evidence from pilot areas in china. *Science of the Total Environment*, 653:565–577.

Feng, Y., Wang, X., Liang, Z., Hu, S., Xie, Y., and Wu, G. (2021). Effects of emission trading system on green total factor productivity in china: Empirical evidence from a quasi-natural experiment. *Journal of Cleaner Production*, 294:126262.

Fowlie, M. (2010). Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review*, 100(3):837–69.

Fowlie, M., Holland, S. P., and Mansur, E. T. (2012). What do emissions markets deliver and to whom? evidence from southern california's nox trading program. *American Economic Review*, 102(2):965–93.

Fu, Y., He, C., and Luo, L. (2021). Does the low-carbon city policy make a difference? empirical evidence of the pilot scheme in china with dea and psm-did. *Ecological Indicators*, 122:107238.

Greenaway, D., Gullstrand, J., and Kneller, R. (2005). Exporting may not always boost firm productivity. *Review of World Economics*, 141(4):561–582.

Greenstone, M. and Gayer, T. (2009). Quasi-experimental and experimental approaches to environmental economics. *Journal of Environmental Economics and Management*, 57(1):21–44.

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C., Garay, M. J., Kalashnikova, O. V., Kahn, R. A., Brauer, M., Apte, J. S., Henze, D. K., Zhang, L., Zhang, Q., Ford, B., Pierce, J. R., and Martin, R. V. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology*, 54(13):7879–7890.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4):605–654.

Hernandez-Cortes, D. and Meng, K. C. (2020). Do environmental markets cause environmental injustice? evidence from california's carbon market. Technical report, National Bureau of Economic Research.

Heyman, F., Sjöholm, F., and Tingvall, P. G. (2007). Is there really a foreign ownership wage premium? evidence from matched employer–employee data. *Journal of International Economics*, 73(2):355–376.

Liu, Y., Tan, X.-J., Yu, Y., and Qi, S.-Z. (2017). Assessment of impacts of hubei pilot emission trading schemes in china–a cge-analysis using termco2 model. *Applied Energy*, 189:762–769.

Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, 105(12):3633–59.

OECD (2008). *OECD Glossary of Statistical Terms*. OECD Publishing.

Qi, S., Cheng, S., and Cui, J. (2021). Environmental and economic effects of china's carbon market pilots: Empirical evidence based on a did model. *Journal of Cleaner Production*, 279:123720.

Ringquist, E. J. (2011). Trading equity for efficiency in environmental protection? environmental justice effects from the so2 allowance trading program. *Social Science Quarterly*, 92(2):297–323.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38.

Shen, J., Tang, P., and Zeng, H. (2020). Does china's carbon emission trading reduce carbon emissions? evidence from listed firms. *Energy for Sustainable Development*, 59:120–129.

Stavins, R. N. (2003). Experience with market-based environmental policy instruments. In *Handbook of environmental economics*, volume 1, pages 355–435. Elsevier.

Tang, K., Zhou, Y., Liang, X., and Zhou, D. (2021). The effectiveness and heterogeneity of carbon emissions trading scheme in china. *Environmental Science and Pollution Research*, 28(14):17306–17318.

van Donkelaar, A., Martin, R. V., Li, C., and Burnett, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, 53(5):2595–2611.

Xu, Y., Jiang, S., Li, R., Zhang, J., Zhao, J., Abbar, S., and González, M. C. (2019). Unraveling environmental justice in ambient pm2. 5 exposure in beijing: A big data approach. *Computers, Environment and Urban Systems*, 75:12–21.

Zhang, W., Li, J., Li, G., and Guo, S. (2020). Emission reduction effect and carbon market efficiency of carbon emissions trading policy in china. *Energy*, 196:117117.

Zhang, W., Zhang, N., and Yu, Y. (2019a). Carbon mitigation effects and potential cost savings from carbon emissions trading in china's regional industry. *Technological Forecasting and Social Change*, 141:1–11.

Zhang, Z., Shao, C., Guan, Y., and Xue, C. (2019b). Socioeconomic factors and regional differences of pm2. 5 health risks in china. *Journal of environmental management*, 251:109564.

Zhou, B., Zhang, C., Song, H., and Wang, Q. (2019). How does emission trading reduce china's carbon intensity? an exploration using a decomposition and difference-in-differences approach. *Science of the total environment*, 676:514–523.

Appendix

T-Test in Nearest-Neighbors Propensity Score Matching

FIGURE 9: OUTCOME OF STATA COMMAND PSTEST FOR NN MATCHING

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
ln_pop	U	6.0328	5.8585	26.2		5.16	0.000	1.12
	M	6.0297	6.0008	4.3	83.4	0.63	0.527	0.99
ln_GDP	U	16.105	15.641	45.1		9.15	0.000	1.31*
	M	16.003	15.972	3.0	93.3	0.44	0.659	0.95
GDP_Per	U	4.1213	2.4604	37.7		10.01	0.000	5.80*
	M	3.2312	3.3115	-1.8	95.2	-0.33	0.745	0.85
secondindustry	U	47.313	49.255	-19.7		-3.49	0.000	0.54*
	M	47.459	47.397	0.6	96.8	0.10	0.923	0.60*
thirdindustry	U	38.865	35.097	47.0		9.20	0.000	1.06
	M	38.071	37.888	2.3	95.1	0.35	0.730	0.75*
ln_FDI	U	10.396	9.1257	70.8		12.90	0.000	0.67*
	M	10.279	10.262	0.9	98.7	0.15	0.881	0.80*
ln_RD	U	11.821	11.443	20.5		3.97	0.000	1.00
	M	11.714	11.685	1.6	92.2	0.23	0.815	0.95
ln_elec	U	13.058	12.506	42.4		8.92	0.000	1.60*
	M	12.931	12.86	5.5	87.1	0.83	0.409	1.35*
afforestation	U	41.721	34.968	27.4		8.06	0.000	12.45*
	M	37.889	38.011	-0.5	98.2	-0.21	0.831	0.92

* if variance ratio outside [0.83; 1.20] for U and [0.83; 1.21] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.115	279.29	0.000	37.4	37.7	74.7*	2.45*	67
Matched	0.002	2.02	0.991	2.3	1.8	9.7	0.87	44

* if B>25%, R outside [0.5; 2]

Time-Varying DID after Matching

TABLE 15: PSM-DID REGRESSION OUTCOMES USING TIME-VARYING ETS TERM

	(1)	(2)	(3)
ETS	-0.103***	-0.104**	-0.112***
	[0.002]	[(0.003]	[0.003]
	(0.0306)	(0.0315)	(0.0338)
City Controls	No	Yes	Yes
Weather Controls	No	No	Yes
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	3943	3740	3347
Adjusted R-squared	0.916	0.935	0.935

Notes : The entries in Table 15 are the coefficient estimates from the PSM-DID model using time-varying ETS terms instead. The city characteristics vector contains population, GDP, GDP per capita, electricity consumption, FDI amount, R&D investment, ratio of secondary industry in GDP, ratio of tertiary industry in GDP, and afforestation rate. The weather control vector contains the annually averaged temperature, humidity, precipitation and sunshine.

Clustered standard errors in parentheses. P-value in square brackets.

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

Map of the Distribution of LCC Regions

FIGURE 10: PARTICIPATING PROVINCES AND CITIES TO THE LCC PROJECT

Source : http://www.tanpaifang.com/

Sensitivity Test Using Alternative Matching Strategies

There may be concerns about the robustness of the PSM-DID approach centered around the choice of propensity score matching strategy. The NN matching process that foster the data for the main DID analysis previously use 4 nearest neighbors to match one treated unit. Utilizing more information to construct the counterfactual for each participant, this method can reduce the variance between the groups at the cost of increased bias. To see if the drawback of the NN matching algorithm consequently corrupts the DID estimation, I thereby adopt other matching algorithms and rerun the PSM-DID procedure. If changing the matching algorithm does not severly reverse the estimation results of PSM-DID model, then the previous estimates for the ETS treatment effect can be viewed as robust.

The three alternative matching algorithms are also commonly used in research, namely the One-to-One matching, Radius matching and Kernel matching. The One-to-One matching finds exactly one comparable untreated unit for each treated unit If either treated or control units are absent in a pairwise block there will be no observation left. Although the comparability between groups will be higher, there is the possibility that some treated units has no available control units to match with and are thus dropped. In contrast, NN matching can find a match for all treated units. The other two matching strategies are also more flexible than One-to-One matching. Radius Matching algorithm assigns each treated unit with only the control units whose propensity score falls into a predefined neighborhood of the propensity score of the treated unit, while in Kernel Matching all treated units are matched with a weighted average of all controls with weights that are inversely proportional to the distance between the propensity scores of treated and controls. For the alternative algorithms to perform better, I follow the suggestions from Austin (2011) to impose caliper/bandwidth equal to 0.2 of the standard deviation of the logit of the propensity score as this value (or one close to it) minimized the mean squared error of the estimated treatment effect in several scenarios.

The PSM-DID regression results under the three alternative matching strategy are displayed in Appendix Table 16. We can see that the CO2 emission reducing effect of ETS is still significant using different matching methods. There is a higher decline in estimate size when applying One-to-One matching as shown in Column (1). Nevertheless, this may be due to the poorer matching performance of this strict matching strategy as shown in Appendix Figure 11 and Figure 14. The number of discarded observations are quite large in this case and the standard biases remain high even after matching. In the case of Radius matching and Kernel matching shown in Column (2) and (3), the results of estimated CO2 emission reduction effects are almost identical to the previous NN matching outcomes. Hence, I argue that the PSM-DID estimations in this paper can endure the changing environment of various matching algorithms.

TABLE 16: DID REGRESSIONS USING ALTERNATIVE MATCHING ALGORITHMS

	(1) One-to-One Matching	(2) Radius Matching	(3) Kernel Matching
ETS	-0.0767*	-0.0786*	-0.0797*
	[0.048]	[0.037]	[0.035]
	(0.0372)	(0.0358)	(0.0361)
Year FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Observations	2065	3730	3738
Adjusted R-squared	0.925	0.935	0.935

Notes : The entries in Table 16 are the coefficient estimates from the PSM-DID regressions using the matched data generated by alternative matching algorithms. Logit regression is used to estimate the propensity score in all three alternative matching strategies as in NN matching. For improving the matching quality, the maximum distance of controls are limited by imposing caliper in One-to-One matching and Radius matching. Equivalently in Kernel matching the bandwidth for kernel is restricted. In DID estimation city characteristics are controlled.

Clustered standard errors in parentheses. P-value in square brackets.

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

TABLE 17: ONE-TO-ONE MATCHING RESULT

Treatment Assignment	Off Support	On Support	Total
Untreated	1818	432	2241
Treated	30	420	450
Total	1848	843	2691

TABLE 18: RADIUS MATCHING RESULT

Treatment Assignment	Off Support	On Support	Total
Untreated	155	2086	2241
Treated	28	422	450
Total	183	2508	2691

TABLE 19: KERNEL MATCHING RESULT

Treatment Assignment	Off Support	On Support	Total
Untreated	152	2089	2241
Treated	23	427	450
Total	175	2516	2691

FIGURE 11: OUTCOME OF STATA COMMAND PSTEST FOR ONE-TO-ONE MATCHING

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
ln_pop	U	6.0328	5.8585	26.2		5.16	0.000	1.12
	M	6.0182	6.0497	-4.7	81.9	-0.47	0.638	0.76*
ln_GDP	U	16.105	15.641	45.1		9.15	0.000	1.31*
	M	15.961	16.158	-19.2	57.5	-1.96	0.050	0.75*
GDP_Per	U	4.1213	2.4604	37.7		10.01	0.000	5.80*
	M	3.056	3.9452	-20.2	46.5	-2.57	0.010	0.52*
secondindustry	U	47.313	49.255	-19.7		-3.49	0.000	0.54*
	M	47.543	48.285	-7.5	61.8	-0.87	0.383	0.61*
thirdindustry	U	38.865	35.097	47.0		9.20	0.000	1.06
	M	37.72	38.814	-13.7	71.0	-1.56	0.120	0.56*
ln_FDI	U	10.396	9.1257	70.8		12.90	0.000	0.67*
	M	10.217	10.793	-32.1	54.7	-3.72	0.000	0.73*
ln_RD	U	11.821	11.443	20.5		3.97	0.000	1.00
	M	11.671	11.73	-3.2	84.4	-0.32	0.747	0.80*
ln_elec	U	13.058	12.506	42.4		8.92	0.000	1.60*
	M	12.875	13.085	-16.2	61.9	-1.62	0.105	1.03
afforestation	U	41.721	34.968	27.4		8.06	0.000	12.45*
	M	37.654	38.451	-3.2	88.2	-1.00	0.319	0.90

* if variance ratio outside [0.83; 1.20] for U and [0.83; 1.21] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.115	279.29	0.000	37.4	37.7	74.7*	2.45*	67
Matched	0.033	20.18	0.017	13.3	13.7	43.6*	0.66	78

* if B>25%, R outside [0.5; 2]

FIGURE 12: OUTCOME OF STATA COMMAND PSTEST FOR RADIUS MATCHING

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
ln_pop	U	6.0328	5.8585	26.2		5.16	0.000	1.12
	M	6.0237	6.0239	-0.0	99.9	-0.00	0.997	1.01
ln_GDP	U	16.105	15.641	45.1		9.15	0.000	1.31*
	M	15.973	15.971	0.2	99.6	0.03	0.976	0.94
GDP_Per	U	4.1213	2.4604	37.7		10.01	0.000	5.80*
	M	3.0851	3.1315	-1.1	97.2	-0.21	0.836	0.91
secondindustry	U	47.313	49.255	-19.7		-3.49	0.000	0.54*
	M	47.529	47.682	-1.6	92.1	-0.24	0.809	0.60*
thirdindustry	U	38.865	35.097	47.0		9.20	0.000	1.06
	M	37.799	37.682	1.5	96.9	0.23	0.819	0.65*
ln_FDI	U	10.396	9.1257	70.8		12.90	0.000	0.67*
	M	10.235	10.226	0.5	99.3	0.08	0.935	0.79*
ln_RD	U	11.821	11.443	20.5		3.97	0.000	1.00
	M	11.685	11.698	-0.7	96.7	-0.10	0.920	0.93
ln_elec	U	13.058	12.506	42.4		8.92	0.000	1.60*
	M	12.89	12.872	1.5	96.6	0.22	0.825	1.33*
afforestation	U	41.721	34.968	27.4		8.06	0.000	12.45*
	M	37.675	37.856	-0.7	97.3	-0.33	0.740	1.01

* if variance ratio outside [0.83; 1.20] for U and [0.83; 1.21] for M

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.115	279.29	0.000	37.4	37.7	74.7*	2.45*	67
Matched	0.001	0.76	1.000	0.9	0.7	6.0	0.85	44

* if B>25%, R outside [0.5; 2]

FIGURE 13: OUTCOME OF STATA COMMAND PSTEST FOR KERNEL MATCHING

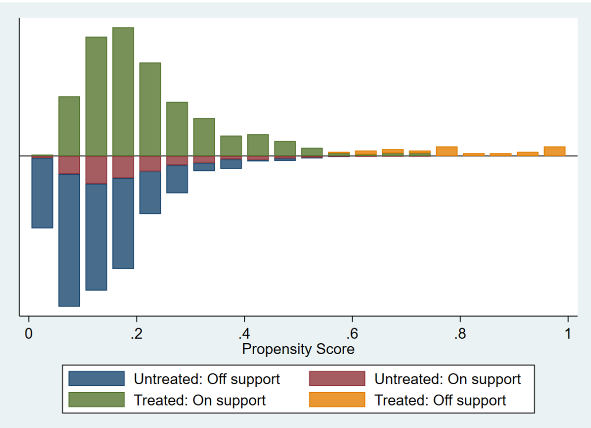
Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct |bias|	t-test t	t-test p>|t|	V(T)/ V(C)
ln_pop	U	6.0328	5.8585	26.2		5.16	0.000	1.12
	M	6.0297	5.9939	5.4	79.4	0.79	0.429	1.03
ln_GDP	U	16.105	15.641	45.1		9.15	0.000	1.31*
	M	16.003	15.948	5.3	88.2	0.79	0.432	0.95
GDP_Per	U	4.1213	2.4604	37.7		10.01	0.000	5.80*
	M	3.2312	3.22	0.3	99.3	0.05	0.963	0.91
secondindustry	U	47.313	49.255	-19.7		-3.49	0.000	0.54*
	M	47.459	47.55	-0.9	95.3	-0.14	0.888	0.59*
thirdindustry	U	38.865	35.097	47.0		9.20	0.000	1.06
	M	38.071	37.51	7.0	85.1	1.05	0.293	0.73*
ln_FDI	U	10.396	9.1257	70.8		12.90	0.000	0.67*
	M	10.279	10.191	4.9	93.1	0.78	0.437	0.76*
ln_RD	U	11.821	11.443	20.5		3.97	0.000	1.00
	M	11.714	11.681	1.8	91.2	0.26	0.792	0.93
ln_elec	U	13.058	12.506	42.4		8.92	0.000	1.60*
	M	12.931	12.84	7.0	83.4	1.07	0.286	1.38*
afforestation	U	41.721	34.968	27.4		8.06	0.000	12.45*
	M	37.889	37.926	-0.1	99.5	-0.06	0.949	0.97

* if variance ratio outside [0.83; 1.20] for U and [0.83; 1.21] for M

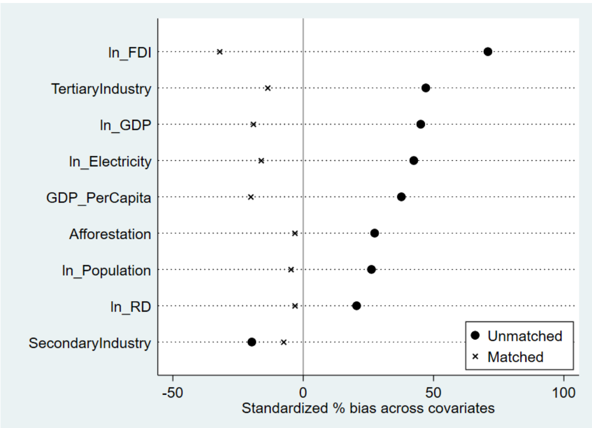
Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%Var
Unmatched	0.115	279.29	0.000	37.4	37.7	74.7*	2.45*	67
Matched	0.002	2.53	0.980	3.6	4.9	10.8	1.21	44

* if B>25%, R outside [0.5; 2]

FIGURE 14: ONE-TO-ONE MATCHING QUALITY

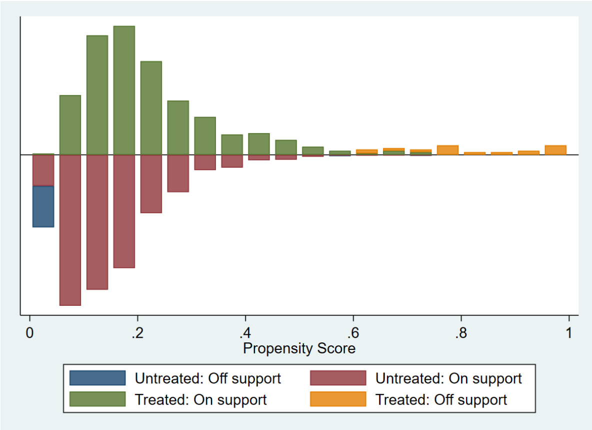


(A) PROPENSITY SCORE HISTOGRAM

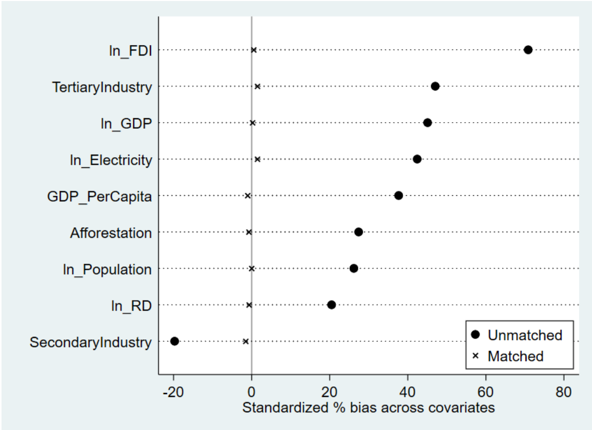


(B) STANDARDISED BIAS BEFORE/AFTER MATCHING

FIGURE 15: RADIUS MATCHING QUALITY

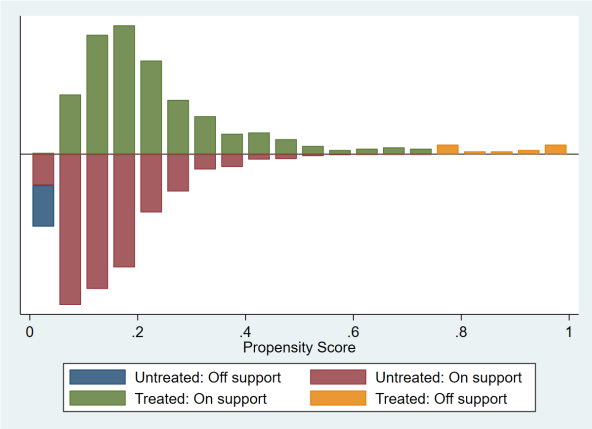


(A) PROPENSITY SCORE HISTOGRAM

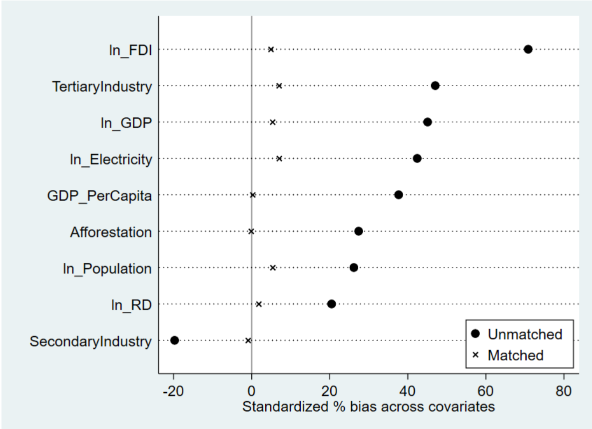


(B) STANDARDISED BIAS BEFORE/AFTER MATCHING

FIGURE 16: KERNEL MATCHING QUALITY

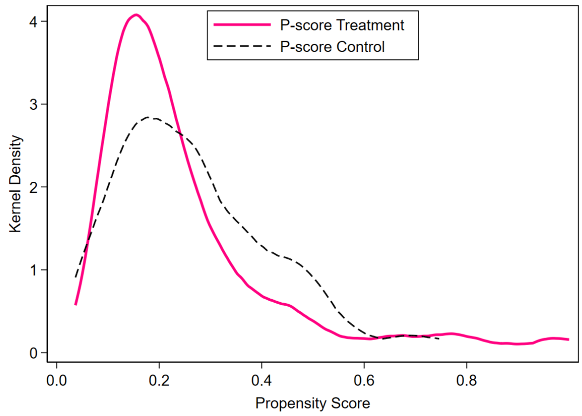


(A) PROPENSITY SCORE HISTOGRAM

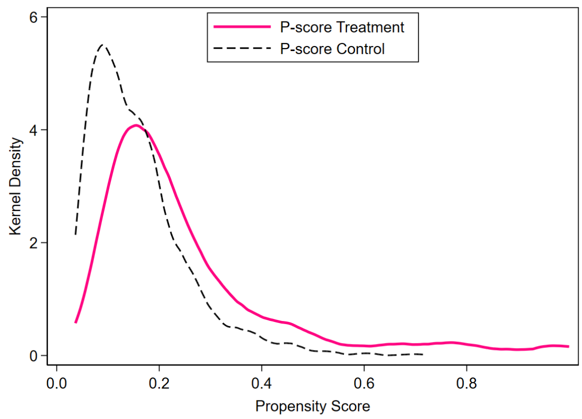


(B) STANDARDISED BIAS BEFORE/AFTER MATCHING

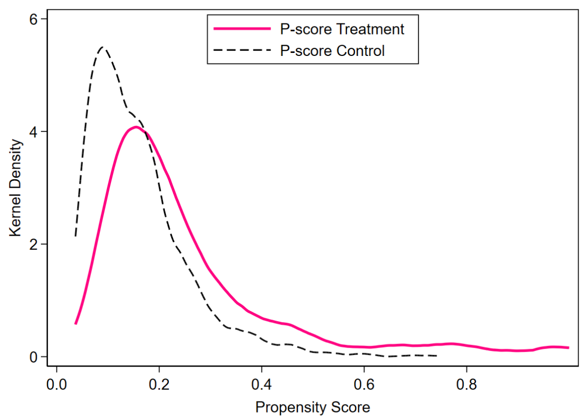
FIGURE 17: MATCHING QUALITY: KERNEL DISTRIBUTION FOR ALTERNATIVE MATCHING



(A) AFTER ONE-TO-ONE MATCHING



(B) AFTER RADIUS MATCHING



(C) AFTER KERNEL MATCHING